

181 GCAAGAGTCAGCATATGATATAATGATTCAGAAATCGTTTGGACGAGTTCGGATGTAGTAG 240
241 TAGCCATTAATTAATGTACATACTAATCGTGAATAGTATGATGAACATCTGATCTT 300
241 TAGCCATTAATTAATGTACATACTAATCGTGAATAGTATGATGAACATCTGATCTT 300
301 ATTGTATAAATATCCATAAACAACATCATGAAGAAGACATTTCTTTCAGGGTCTGAATTAAT 360
301 ATTGTATAAATATCCATAAACAACATCATGAAGAAGACATTTCTTTCAGGGTCTGAATTAAT 360
361 TATGATACAAATCTTAATAGAAAAACGAATTAATTAAGTTGAATTTGATGAATCTAAATTG 420
361 TATGATACAAATCTTAATAGAAAAACGAATTAATTAAGTTGAATTTGATGAATCTAAATTG 420
421 AACAGCCAAACACGACGAGGACTAATCGTTGCTGGATGACTCGGTTTAAGTTAACCCAC 480
421 AACAGCCAAACACGACGAGGACTAATCGTTGCTGGATGACTCGGTTTAAGTTAACCCAC 480
481 TAAAAAAACGGAGCTGTCATGTAAACACGCGGATCGAGCAGGTACAGTCAATGAAGCCATC 540
481 TAAAAAAACGGAGCTGTCATGTAAACACGCGGATCGAGCAGGTACAGTCAATGAAGCCATC 540
541 AAAGCAAAAGAACTAATCCAGGGGTGAGATTAATTAATTAAGTTTAAAAATTAAGTTAACAC 600
541 AAAGCAAAAGAACTAATCCAGGGGTGAGATTAATTAATTAAGTTTAAAAATTAAGTTAACAC 600
601 GAGGGAAGAGCTGTCGACAGCAGGTACGTTATCTTTACCTGTGTGTCGAAATGATTCG 660
601 GAGGGAAGAGCTGTCGACAGCAGGTACGTTATCTTTACCTGTGTGTCGAAATGATTCG 660
661 TGTCTGTCCGATTTAAATTAATTTTGAAGGCCGAAATAAGTTGTAAGAGATAAACC 720
661 TGTCTGTCCGATTTAAATTAATTTTGAAGGCCGAAATAAGTTGTAAGAGATAAACC 720
721 CGCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTGTGTTCTCTCTCAC 780
721 CGCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTGTGTTCTCTCTCAC 780
781 TTTTCATCAGCGTTTGAATCTCCGCGACTTGAACAGAGAAGAACAGAGAAGACTAA 840
781 TTTTCATCAGCGTTTGAATCTCCGCGACTTGAACAGAGAAGAACAGAGAAGACTAA 840
841 GAGAGAACTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTTCCTCAATCTC 900
841 GAGAGAACTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTTCCTCAATCTC 900
901 ATCTTCTTCTCCGCTCTTTCTTCCAGGTAATAGGAATCTTCTGGATCTACTTTATTT 960
901 ATCTTCTTCTCCGCTCTTTCTTCCAGGTAATAGGAATCTTCTGGATCTACTTTATTT 960
961 GCTGGATCTGATCTGTTTCTCAATTTCTTCCGCTTGAATCTGGAATCTGTTTAAATTGGAT 1020
961 GCTGGATCTGATCTGTTTCTCAATTTCTTCCGCTTGAATCTGGAATCTGTTTAAATTGGAT 1020
1021 CTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGATCTAAGTTGACCGATCAGT 1080
1021 CTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGATCTAAGTTGACCGATCAGT 1080
1081 TAGCTCGAATATAGCTTACAGAAATTTGGCTTGAACCTTGAATGAGAGATCCATGTTCAATG 1140
1081 TAGCTCGAATATAGCTTACAGAAATTTGGCTTGAACCTTGAATGAGAGATCCATGTTCAATG 1140
1141 TACCTGGGAAATGATTTGATATATGTAATCTGAATCTGAATCTGGAATCTGGAATGATTAAT 1200
1141 TACCTGGGAAATGATTTGATATATGTAATCTGAATCTGAATCTGGAATCTGGAATGATTAAT 1200
1201 CTGAACACTGTAATCTGATGTAATCTGAACACTGTTTAACTTGAATGAGAGATTTGCTG 1260
1201 CTGAACACTGTAATCTGATGTAATCTGAACACTGTTTAACTTGAATGAGAGATTTGCTG 1260
1261 ATAGATTTCTTCAAAACCTTAGGATTTGATGTGTGTCGATCGTGAACAGAAAGCTATTTCTG 1320

1261 ATAGATTTCTTCAAAACCTTAGGATTTGATGTGTGTCGATCGTGAACAGAAAGCTATTTCTG 1320
1321 ATTCAATCAGGTTTATTTGATGATGATGAACTCTTTTGTGTGTTTGCAGCTCATAAAA 1380
1321 ATTCAATCAGGTTTATTTGATGATGATGAACTCTTTTGTGTGTTTGCAGCTCATAAAA 1380
1381 AATGGCTGAGGCTGACGATATTTCAACCAATCGTGTGTGACAAATGGTACTTGGAAATGGTAGG 1440
1381 AATGGCTGAGGCTGACGATATTTCAACCAATCGTGTGTGACAAATGGTACTTGGAAATGGTAGG 1440
1441 ATCC 1444
1441 ATCC 1444

RESULT 2
AB026654
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB026654 75508 bp DNA linear PLN 14-FEB-2004
Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone; MVE11.
AB026654 BA000014
AB026654.1 GI:4757410
Arabidopsis thaliana (chale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS
TITLE

Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S.
Structural analysis of Arabidopsis thaliana chromosome 3. I.
Sequence features of the regions of 4,504,864 bp covered by sixty
P1 and TAC clones
DNA Res. 7 (2), 131-135 (2000)
20277480
10819329

JOURNAL
MEDLINE
PUBMED

2 (bases 1 to 75508)
Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S.
Direct Submission
Submitted (28-APR-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MVE11
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://grem1inl.2001.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is K24M9 and the 3' clone is MCB22.

FEATURES
source
1..75508
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MVE11"
/clone_lib="Mitsui P1"
/ecotype="Columbia"
/note="unnamed protein product; gene_id:MVE11.1"

CDS

unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01791.1"
/db_xref="GI:9293888"
/translation="MASSSSSSSSTSKKNTFKKNSAFSSSRATSLIKRQQTVP
KAKELSLICDDIVCVIGSNGELKTWEEREKVAIARRIGELSETKRRGSVDLH
EFLEKNNDDPEKEKKIKIVRRPVKVPVWDPRFDNYVSELQMLVQSLERNLRI
QRTCAVBAQOORVQVYTNMANQELMANTNQLQOHSNQSYMILNHNHNGAFSQIP
VSALASNOTQS LAPPELMYIPNSDAGNYSGSGVQGTINGLQNMNMLTYNNINSV
NDFSQFQONSRAESYSLLGVHEDGNNEFPNMSRNNFVQDCAGLLGQAGATN
GLQSMWMDYSNNSNLSHGVQVQPTYSQHQRVFNLDONGNTRSL"
complement (join(3045..3317,3774..4546,4652..5271,
5392..5628,5717..5793))
/note="unnamed protein product; gb|AAC34345.1
gene_id:MVE11.2
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01792.1"
/db_xref="GI:9293889"
/translation="MANGPAAPAPTITTTGGDSRRRLSASIEAICRRFRNRKSGGGRS
DMVKPFNIINFTQDKNSCCCTKFQIVKLLILLATLPTLIYSPVAHSHSHS
SRNLTWRQDPYRFDLDINDDVTKLENIIEGRTIVLNFDNSNEIQRWREVSXKD
NGDEKVVNLNDYADKNVNDALYPEWIDEEQTEVPVCPNIPNKTRKRLDIVV
KLPCKEGNSRDVGRHLQLAAATAAASAKGFFRGVFFVSRCPFINLFRCDLVS
RRGDVLYKPNLDRDLKQLPVGSELSPLGIQDRPSLGNPKREAYATILHSADV
VCAAAQASIRQSGSTRDLVTLVDNITSGYHRSGLAAGWOIRTIQRNPKASKDA
YCNWYSFRLWLTDXKIFIDADLLILRNI DFLFMPETISATGNGTLPNSGMV
LEPCNCTQLLMEHNEIESYNGGQGVLYNEVTFWHRIIPKHNFLKHFVIGDEDDAK
RKTLFQAGPEPVLVLYHGLMKWCYRDYDCNFNSDIFVEFATDIARRKNNWVHDA
MQEULQCYRKSQAQLEYYDRROAEAAAYADGHWKIRVKDPRFKICIDKLCNNKSM
LRHWGESNWDYVESVTPPAITVDRRSSLPGHNL"
complement (join(7969..8583,8684..9123,9207..9332,
9437..9586,9675..10305))
/note="unnamed protein product; emb|CAB70981.1
gene_id:MVE11.3
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01793.1"
/db_xref="GI:9293890"
/translation="MDTEKGYPRGFRASLRPETGPRFRFDSMQIDBAADLQKDIEIR
QENSYLVFNKIDSGIELEATQFLDRNPEALTALTITNGDTPHKAVALSGHIVKEE
IIRIHDPEQVLKINDNGYALTAYATGAGIVIAECLVKNCPGIVSVRNAKEHIPV
VARSYGHKHLVOYLYSHTPSLDPCDDSDHEKGNKAMLVNCTIVDGLYCTALDLIQ
RYPKLAYTRSDNDTAIMALACTPYAFPSGTETAFQWRWYVSCIHIEKINNPHEVNH
HHHQYRNSQDHSQEBKMHKLYLFRIRIRVYVYKLGHAQAKEILDICQELPKF
DAAQQNAGLQALPQAVENGIVEYIEEMRHYPDI VWSKNSSNGNIIFYAVSQRQKE
IFSLYNTIGAKNNILATNDIPHNMLHHAAYRAPASRLNIPGAALQMQRELQWPK
VEKLQPKHRKVNVLKQKTPALFDQHKDILVEQEKWKETATCTVVAALITTM
FSSAFTVPGYRSDCMPLYIHOHRFKIFLISDAISLFTSCMSLLMFLGILKSRVRED
FLRSIPTKLI VGLLALFLSMATMIVTVVLTMTLVGEKISVWSAQFMLAVIPGMFV
VLQPVVLEIIPATYCPNVFQKPRVFKL"
join(11367..11581,11799..11940,12035..12088,12453..12544,
12656..12882,12966..13142,13242..13354)
/note="unnamed protein product; contains similarity to
uridylylate kinase
gene_id:MVE11.4"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01794.1"
/db_xref="GI:9293891"
/translation="MAIPLTSCSPISTSSISRTSVPLTRNRTFFSNVSRVV
LISCSLSNGSSGSDPMNGNGSLNGQSDFPLPSPDGTSGKPLNWRVLLK
VSGEALADGEONIDPKVIMALAREVAATRLGIEVALVPGGNI FRGSTWAGCSGLD
RSSADYIGMLATVMAIPLQATMESIGIPTRVQAFRMSVABPYIRRAIRHLEKGR
WVFAAGTGNPFTTDTAALRCABINAEVVLKATNVGDFDDPFRPNRNLRLSLT
YQVTSKDLSDVMDMTAILTCQENNI PVVVFNLSEPCNTAKAIKGERVCTLIGGTWNSI
VTTS"
complement (13589..14257)
/note="unnamed protein product; gb|AAF16598.1
gene_id:MVE11.5

similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01795.1"
/db_xref="GI:9293892"
/translation="MDPSEYFAGGNPSDOONKROQLQICGPRPSPLSVHKDSHKIKKP
PKHPAPPNRQDPPYIPREPVIYAVSPKVHATASEFMNVVVGITGSSGVFLES
GGDVSPPAARLASTENASPRGKEPAARDETVEINTAMEAAAFQGYAGFGLSPSPAL
LPTASTGTFSPMYHOGGMFSPAIPGLGLFPAGFMSPPRSGTSLVASPTFADFSTHI
WDQST"
complement (join(15556..15731,16089..16154,16246..16333,
16599..16739))
/note="unnamed protein product; gene_id:MVE11.6
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01796.1"
/db_xref="GI:9293893"
/translation="MPTSYTSPSAOPTNNWFLFNIFSSLLVIALRFFDQDVSDRY
SYKONGLFLNENAFVKKQIGLKFNKIITNMKBKAVLVENEIGLKFNKI KPVKGBEY
PCPSLPALPDTKASSNSDLDMAPNHEYQVSVNSPTLLANDSTGEVSKGPCWSHLL"
complement (18068..19135)
/note="unnamed protein product; gene_id:MVE11.7
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01797.1"
/db_xref="GI:9293894"
/translation="MGRDETETITVPSPFKCIPISLDVMSRSPVSLCTGVYDRAIQ
WIDGNNTCPATMQLLTKTDFPNLTQLRLINWSDSIGRRHNGSPVLNPPSGREVP
TKBEYNLLERLSLENLMKIVRVKVDSDSNREFSKWBFVPMVLVDIIRTKTKIEL
VIMAIRYIKVDRERLNLMLANDGGDCLTALLAIQORGLNESIESRVLDWISF
DAKSLMIAETRGVLTEMMKTSITESSDPSLIEASLSPLITISKSRRVSLIAAKA
ITKQILTLTNNVAVTEKSLKLETLSSKREGRLEICGDNRCRVEGVVKKLLKV
STTAHAATVILMCLCYVEREDKTVEETVERSNGVTKLLVLIQSNCSAMVRQAKDLI
KVLNPSALLAAEYKTHIMP"
complement (join(28267..29317,29494..29585))
/note="unnamed protein product; gb|AAD43172.1
gene_id:MVE11.8
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01798.1"
/db_xref="GI:9293895"
/translation="WTRRTQYPIAELSLARSIVQALPASATISGNGPKKKKNCV
RGLMDKQIPTLLQELIIRGLKANIHASLCKTWLKEAVSRKSPMLPYPOSQR
GGPEGDVYLNPSRQTHLKFELTCYRNKLACAKDGLLVKXNDPVDVFLNPF
GERICLPQPNSTRDCLTFSAAPTSCCVISFTPSQFLYAVVKDVTWRPGESVMT
HFQPKRYGEVNRICIFSNMGFYCLSTSGRLSVDFPSRETNVLPKPCRAFRKIML
VRQVMEHEGDI FVVTTRVNRKLLAFKLNQNVWEEKVPMKPLTVFSSDATSLT
RAGLPEERNLISDIDDFVKSHPTFYDYDCSAWLOPHDNFNF"
complement (join(29896..30389,30465..30582,30745..30879,
31009..31129,31210..31276,31364..31468,31677..31788,
31869..31981,32106..32223,32374..32474,32694..32810,
33023..33152,33271..34000,34112..34210,34701..34792,
35000..35176,35688..36488,36613..36879))
/note="unnamed protein product; gene_id:MVE11.9
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01799.1"
/db_xref="GI:9293896"
/translation="MGRDVAARAKYAKAEVGRDRRQARWANNVDILKNHGYVD
ALKWFRIDYDLSVKYLPQKDLPTCOSIGETVLRLENFEALIVYKHQLAERANDT
VEKQACQTLQRTTHEMPLKEDCEALQSAKYPKAMELAQILKEKPPGESSGFL
EYINAHNNI GMLDLDNPEARTILKKGLQICDEEVRDYDAARSRHHNLGNVFM
ALRMSWDKSALEQI EHNKTIVKSKMWKEELREBELMLKCLSAEMTDKGTSEERKSM
AKSQDSSALVEQI EHNKTIVKSKMWKEELREBELMLKCLSAEMTDKGTSEERKSM
LOVNACLSLIDKSSVMVFLQYSKRKKISDELCDKEKLDAPFMVGESYONLR
NFRSLKWFIRSYESGHEAIGNLEQALAKINIGLDCICIGETWGLQAYEGRYALK
ANLPSIQSALIEDIHYHMRFGNAQKASELKEIQLNKESEHAKEKSCSTQDECS

Query Match

93.5%; Score 1349.6; DB 8; Length 75508;

Best Local Similarity 98.7%; Pred. No. 2.8e-281; Matches 1423; Conservative 0; Mismatches 11; Indels 8; Gaps 6;																																					
QY	1	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATATATGTAAAG	60																																		
Db	57986	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATATATGTAAAG	58045																																		
QY	61	AAAGTTTTCACGAATATGNNACGACAAAATGGCTACACTCGATGTAAATGGGTATCTCAAC	120																																		
Db	58046	AAAGTTTTCACGAATATGGCAGCAGAAAATGGCTAGACTCGATGTAAATGGGTATCTCAAC	58105																																		
QY	121	TCACATTTATCTATPACCAAAACATTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATGT	178																																		
Db	58106	TCACATTTATCTATPACCAAAACATTAGTTAGCAAAAATTTAAACAACCTATTTTATGT	58165																																		
QY	179	ATGCAAGAGTCAGCATATGTATAATTTGATTCGAATCGTTTTGACGAGTTCGGATGTAGT	238																																		
Db	58166	ATGCAAGAGTCAGCATATGTATAATTTGATTCGAATCGTTTTGACGAGTTCGGATGTAGT	58225																																		
QY	239	AGTAGCCATTATTTAATGTACATACATAATCGTGAATAGTG-ATATGATGAAACATTGTAT	297																																		
Db	58226	AGTAGCCATTATTTAATGTACATACATACTAATCGTGAATAGTGAATATGATGAAACATTGTAT	58285																																		
QY	298	CTTATTGTATAAATPCCATAAACACATCATGAAAGACACTTTCTTTCAAGGCTCTGAATT	357																																		
Db	58286	CTTATTGTATAAATPCCATAAACACATCATGAAAGACACTTTCTTTCAAGGCTCTGAATT	58345																																		
QY	358	AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGTATGAAATCTTAA	417																																		
Db	58346	AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGTATGAAATCTTAA	58405																																		
QY	418	TTGAACAAGCCAACACGACGAGGACTAACGTTGCTGATTTGACCTCGGTTTAAAGTTAAC	477																																		
Db	58406	TTGAACAAGCCAACACGACGAGGACTAACGTTGCTGATTTGACCTCGGTTTAAAGTTAAC	58465																																		
QY	478	CACATAAAAAACGGAGCTGTATGTAAACACGCGGATCGAGCAGGTACAGTCTATGAAGCC	537																																		
Db	58466	CACATAAAAAACGGAGCTGTATGTAAACACGCGGATCGAGCAGGTACAGTCTATGAAGCC	58525																																		
QY	538	ATCAAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	597																																		
Db	58526	ATCAAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	58585																																		
QY	598	CACAGGGGAAAA-GCTGTCTGTACAGCCAGGTACAGTTATCTTTACCTGTGGTCGAAATCA	656																																		
Db	58586	CACAGGGGAAAAAGGCTGTGTACAGCCAGGTACAGTTATCTTTACCTGTGGTCGAAATCA	58645																																		
QY	657	TTCTGTCTGTCAATTTTAAATTTTGAAGGCCGAAAAATAAAGTTGTAAAGAGATA	716																																		
Db	58646	TTCTGTCTGTCAATTTTAAATTTTGAAGGCCGAAAAATAAAGTTGTAAAGAGATA	58705																																		
QY	717	AACCCGCCCTATAAATTCATATATTTTCTCCCGCCTTTGAAATGTCTCGTTGCTCTCC	776																																		
Db	58706	AACCCGCCCTATAAATTCATATATTTTCTCTCCCGCCTTTGAAATGTCTCGTTGCTCTCC	58765																																		
QY	777	TCACCTTTTCATCAGCCGTTTTTGAATCTCCGCGACTTTGACAGAGAAACAAGGAAGA	836																																		
Db	58766	TCACCTTTTCATCAGCCGTTTTTGAATCTCCGCGACTTTGACAGAGAAACAAGGAAGA	58825																																		
QY	837	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTGTGAATCTTCTCTAA	896																																		
Db	58826	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTGTGAATCTTCTCTAA	58885																																		
QY	897	TCATCTCTCTCTCCGCTCTTTCTTTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT	956																																		
Db	58886	TCATCTCTCTCTCTCCGCTCTTTCTTTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT	58942																																		
QY	957	ATTTCGTGATCTCGATCTGTTTCTCAATTTCTTTGAGATCTGGAATTCGTTTAAATTT	1016																																		
Db	58943	ATTTCGTGATCTCGATCTGTTTCTCAATTTCTTTGAGATCTGGAATTCGTTTAAATTT	59002																																		
QY	1017	GGATCTGTGAACCTCCATAAATCTTTTGGTTTTTACTAGAAATCGATCTCAAGTTGACCGAT	1076																																		
Db	59003	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAAATCGATCTAAGTTGACCGAT	59062																																		
QY	1077	CAGTTAGCTCGAATTATAGCTACAGAAATTTGGCTTTGACCTTCATGGAGAGATCCATGTTC	1136																																		
Db	59063	CAGTTAGCTCGAATTATAGCTACAGAAATTTGGCTTTGACCTTCATGGAGAGATCCATGTTC	59122																																		
QY	1137	ATGTTACCTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAACTGTTTGAAGTTAGATT	1196																																		
Db	59123	ATGTTACCTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAACTGTTTGAAGTTAGATT	59182																																		
QY	1197	GAACTGAAACACTGTCAAATGTTAGATTGAACTGAACTGAACTGTTTGAAGTTAGATT	1255																																		
Db	59183	GAACTGAAACACTGTCAAATGTTAGATTGAACTGAACTGTTTGAAGTTAGATT	59242																																		
QY	1256	TGTTGATAGATTCTTCGAAACCTTAGGATTTCTAGTCTGTTGAACTGTTGAACTGTTGAACTGTT	1315																																		
Db	59243	TGTTGATAGATTCTTCGAAACCTTAGGATTTCTAGTCTGTTGAACTGTTGAACTGTTGAACTGTT	59302																																		
QY	1316	TTCTGATTCAAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTTTGTGTTTGCAGCTCA	1375																																		
Db	59303	TTCTGATTCAAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTTTGTGTTTGCAGCTCA	59362																																		
QY	1376	TAAAAAATGGCTGAGGCTGACGATATTCAACCAATCGTGTGTGACAACTGGTACTGGAATG	1435																																		
Db	59363	TAAAAAATGGCTGAGGCTGATGATATTCAACCAATCGTGTGTGACAACTGGTACTGGAATG	59422																																		
QY	1436	GT 1437																																			
Db	59423	GT 59424																																			
RESULT 3																																					
AX463618																																					
LOCUS	AX463618	6385 bp	DNA	linear	PAT 15-JUL-2002																																
DEFINITION	Sequence 15 from Patent W00248335.																																				
ACCESSION	AX463618																																				
VERSION	AX463618.1	GI:21886378																																			
KEYWORDS	synthetic construct																																				
SOURCE	synthetic construct																																				
ORGANISM	other sequences; artificial sequences.																																				
REFERENCE	1																																				
AUTHORS	Meagher, R.B. and Li, Y.																																				
TITLE	Metal resistant plants and phytoremediation of environmental																																				
JOURNAL	contamination																																				
Patent: WO 0248335-A 15 20-JUN-2002;																																					
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)																																					
FEATURES																																					
source																																					
1. .6385																																					
/organism="synthetic construct"																																					
/mol_type="unassigned DNA"																																					
/db_xref="taxon:32630"																																					
/note="Plasmid pACT2B-GUS"																																					
ORIGIN																																					
Query Match 92.9%; Score 1341.8; DB 6; Length 6385;																																					
Best Local Similarity 98.8%; Pred. No. 1.8e-279;																																					
Matches 1375; Conservative 0; Mismatches 12; Indels 4; Gaps 2;																																					
QY	1	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATATATGTAAAG	60																																		
Db	257	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATATATGTAAAG	316																																		
QY	61	AAAGTTTTCACGAATATGNNACGACAAAATGGCTACACTCGATGTAAATGGGTATCTCAAC	120																																		
Db	317	AAAGTTTTCACGAATATGNNACGACAAAATGGCTACACTCGATGTAAATGGGTATCTCAAC	376																																		
QY	121	TCACATTTATCTATPACCAAAACATTAGTTAGCAAAAATTTAAACAACCTATTTTATGTAT	180																																		
Db	377	TCACATTTATCTATPACCAAAACATTAGTTAGCAAAAATTTAAACAACCTATTTTATGTAT	436																																		
QY	181	GCAAGAGTCAGCATATGTATATAATTTGATTCAGAAATCGTTTTGACGAGTTCGGATGTAGTAG	240																																		

```
Db 437 GCAGAGTCCAGCATATGTATAATTGATTGAGTTCAGAAATCGTTTTTGACGAGTTCGGATGTAGTAG 496
Qy 241 TAGCCATTATTTAATGTACATACTAATCGTGAATAGTATGATGATGAACATTTGTATCTTT 300
Db 497 TAGCCATTATTTAATGTACATACTAATCGTGAATAGTATGATGATGAACATTTGTATCTTT 556
Qy 301 ATTGTATAAATATCCATAACACATCATGAAGACACTTTCTTTGAGGTCGTAATTAAT 360
Db 557 ATTGTATAAATATCCATAACACATCATGAAGACACTTTCTTTGAGGTCGTAATTAAT 616
Qy 361 TATGATACAAATTTAATAGAAACGAATTAATTTAGTTGAAATTTGATGAAATCTAAATTG 420
Db 617 TATGATACAAATTTAATAGAAACGAATTAATTTAGTTGAAATTTGATGAAATCTAAATTG 676
Qy 421 AACAGCCCAACACGACGAGGACTAACCGTTGCTGGAATGACTCGGTTTAAAGTTAAACAC 480
Db 677 AACAGCCCAACACGACGAGGACTAACCGTTGCTGGAATGACTCGGTTTAAAGTTAAACAC 736
Qy 481 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTCAAGTCATGAAGCCATC 540
Db 737 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTCAAGTCATGAAGCCATC 796
Qy 541 AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAAATTTAGTTTAAAAATTTAGTTAAACAC 600
Db 797 AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAAATTTAGTTTAAAAATTTAGTTAAACAC 856
Qy 601 GAGGAAAAA - GCTGTCTGACAGCAGGTCAAGTTATCTTTTACCTGTGTGCGAAATGATTC 659
Db 857 GAGGAAAAAGGCTGCTGACAGCCAGGTCAAGTTATCTTTTACCTGTGTGCGAAATGATTC 916
Qy 660 GTGTCTGTGCAATTTAATTTATTTTAAAGGCGGAAATAAAGTTGTAAGAGATAAAC 719
Db 917 GTGTCTGTGCAATTTAATTTATTTTAAAGGCGGAAATAAAGTTGTAAGAGATAAAC 976
Qy 720 CGGCTATATAAATTCATATATTTTCCCTCCCGCTTTGAAATTTGCTCGTTGCTCCCTCA 779
Db 977 CGGCTATATAAATTCATATATTTTCCCTCCCGCTTTGAAATTTGCTCGTTGCTCCCTCA 1036
Qy 780 CTTTCATCAGCGCTTTTGAATCTCGGCGACTTGACAGAGAAAGAAAGAAAGAACTA 839
Db 1037 CTTTCATCAGCGCTTTTGAATCTCGGCGACTTGACAGAGAAAGAAAGAAAGAACTA 1096
Qy 840 AGAGAGAAAGTAAAGAGATAATCCAGAGATTCATTTCTCGGTTTGAATCTTCTCAATCT 899
Db 1097 AGAGAGAAAGTAAAGAGATAATCCAGAGATTCATTTCTCGGTTTGAATCTTCTCAATCT 1156
Qy 900 CATCTCTCTTCTCGCTCTTCTTCCAGAGTAATAGGAACCTTTCTGGATCTACTTTAT 959
Db 1157 CA---TCCTTCTTCTCGCTCTTCTTCCAGAGTAATAGGAACCTTTCTGGATCTACTTTAT 1213
Qy 960 TCCTGGATCTCGATCTTGTCTTCTCAATTTCTCTTGAGATCTGGAATTCGTTTAAATTTGA 1019
Db 1214 TCCTGGATCTCGATCTTGTCTTCTCAATTTCTCTTGAGATCTGGAATTCGTTTAAATTTGA 1273
Qy 1020 TCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAATCGAATGACCGATCAG 1079
Db 1274 TCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAATCGAATGACCGATCAG 1333
Qy 1080 TTAGCTCGAATTAAGTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCAGTTTCATG 1139
Db 1334 TTAGCTCGAATTAAGTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCAGTTTCATG 1393
Qy 1140 TTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACTGTGTAAGTTAGATTTGAA 1199
Db 1394 TTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACTGTGTAAGTTAGATTTGAA 1453
Qy 1200 TCTGAACACTGTCAATGTGTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTG 1259
Db 1454 TCTGAACACTGTCAATGTGTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTG 1513
Qy 1260 TATAGATTTCTGAAACCTTAGGATTTGTAGTGTGTAGTGTGTAACAGAGAAAGCTATTCT 1319
Db 1514 TATAGATTTCTGAAACCTTTAGGATTTGTAGTGTGTAGTGTGTAACAGAGAAAGCTATTCT 1573
```

```
Qy 1320 GATTCAATCAGGGTTTATTTGACTGTATTTGAACCTCTTTTGTGTGTTTGCAGTCTATAAA 1379
Db 1574 GATTCAATCAGGGTTTATTTGACTGTATTTGAACCTCTTTTGTGTGTTTGCAGTCTATAAA 1633
Qy 1380 AAATGGCTGAG 1390
Db 1634 GGATCCCCGGG 1644

RESULT 4
AX463617
LOCUS AX463617 4526 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 14 from Patent WO0248335.
ACCESSION AX463617
VERSION AX463617.1 GI:21886377
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Meagher,R.B. and Li,Y.
TITLE Metal resistant plants and phytoremediation of environmental
JOURNAL contamination
Patent: WO 0248335-A 14 20-JUN-2002;
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)
FEATURES
source
1. 4526
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pACT2B"

ORIGIN
Query Match 92.9%; Score 1341.4; DB 6; Length 4526;
Best Local Similarity 98.9%; Pred. No. 2.3e-279;
Matches 1374; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

Qy 1 ATTATGATCTCAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATTATGTAG 60
Db 257 ATTATGATCTCAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATTATGTAG 316
Qy 61 AAAGTTTGAAGAAATATGNNACGACAAATGGCTACACTCGATGTAAATGGGTATCTCAAC 120
Db 317 AAAGTTTGAAGAAATATGNNACGACAAATGGCTACACTCGATGTAAATGGGTATCTCAAC 376
Qy 121 TCACATTTATATCTTATACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTATGTAT 180
Db 377 TCACATTTATATCTTATACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTATGTAT 436
Qy 181 GCAAGAGTCAGCATATGTATAATTGATTTCAGATTCGATTCGAGATTCGGATGTAGTAG 240
Db 437 GCAAGAGTCAGCATATGTATAATTGATTTCAGATTCGATTCGAGATTCGGATGTAGTAG 496
Qy 241 TAGCCATTATTTAATGTACATACTAATCGTGAATAGTATGATGATGAACATTTGTATCTTT 300
Db 497 TAGCCATTATTTAATGTACATACTAATCGTGAATAGTATGATGATGAACATTTGTATCTTT 556
Qy 301 ATTGTATAAATATCCATAACACATCATGAAGACACTTTCTTTGAGGTCGTAATTAAT 360
Db 557 ATTGTATAAATATCCATAACACATCATGAAGACACTTTCTTTGAGGTCGTAATTAAT 616
Qy 361 TATGATACAAATTTAATAGAAACGAATTTAAATTTAGTTGAAATTTGATGAAATCTAAATTG 420
Db 617 TATGATACAAATTTAATAGAAACGAATTTAAATTTAGTTGAAATTTGATGAAATCTAAATTG 676
Qy 421 AACAGCCCAACACGACGAGGACTAACCGTTGCTGGAATGACTCGGTTTAAAGTTAAACAC 480
Db 677 AACAGCCCAACACGACGAGGACTAACCGTTGCTGGAATGACTCGGTTTAAAGTTAAACAC 736
Qy 481 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTCAAGTCATGAAGCCATC 540
Db 737 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTCAAGTCATGAAGCCATC 796
```

Qy	541	AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAAACAC	600		
Db	797	AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAAACAC	856		
Qy	601	GAGGGAAGA--GCTGTCTGTCAGCAGCGTCACTTTACTTTACCTGTGGTCGAAATGATTC	659		
Db	857	GAGGGAAGAAGCGTGTCTGACAGCAGGTCACTTTACTTTACCTGTGGTCGAAATGATTC	916		
Qy	660	GTGTCTCTCGATTTTAATTAATTTTTTGAAGGCCGAAAAATAAGTTGTAAGAGATAAAC	719		
Db	917	GTGTCTCTCGATTTTAATTAATTTTTTGAAGGCCGAAAAATAAGTTGTAAGAGATAAAC	976		
Qy	720	CCGCTATATAAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCTCA	779		
Db	977	CCGCTATATAAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCTCA	1036		
Qy	780	CTTTTCATCAGCGCTTTTGAATCTCCGCGGACTTGACAGAGAACAAGAAAGAGACTA	839		
Db	1037	CTTTTCATCAGCGCTTTTGAATCTCCGCGGACTTGACAGAGAACAAGAAAGAGACTA	1096		
Qy	840	AGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCT	899		
Db	1097	AGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCT	1156		
Qy	900	CATCTTCTTCTCCGCTCTTTCTTTTCCAAAGTAATAGGAACCTTTCTGGAATCTACTTTATT	959		
Db	1157	CA--TCTTCTCCGCTCTTTCTTTTCCAAAGTAATAGGAACCTTTCTGGAATCTACTTTATT	1213		
Qy	960	TGCTGGATCTCGATCTTGTTTTCTCAATTTCTTGAGATCTGGAATTCGTTTAAATTTGGA	1019		
Db	1214	TGCTGGATCTCGATCTTGTTTTCTCAATTTCTTGAGATCTGGAATTCGTTTAAATTTGGA	1273		
Qy	1020	TCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAACTGAATCTGAATCTGATCTG	1079		
Db	1274	TCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAACTGAATCTGAATCTGATCTG	1333		
Qy	1080	TTAGCTCGAATATAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCATGTTTCATG	1139		
Db	1334	TTAGCTCGAATATAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCATGTTTCATG	1393		
Qy	1140	TTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCTGTTGAAAGTTAGATTGAA	1199		
Db	1394	TTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCTGTTGAAAGTTAGATTGAA	1453		
Qy	1200	TCTGAACACTGTCGAATCTAGATTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTG	1259		
Db	1454	TCTGAACACTGTCGAATCTAGATTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTG	1513		
Qy	1260	TATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGACGTTGAACAGAAAGCTATTTC	1319		
Db	1514	TATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGACGTTGAACAGAAAGCTATTTC	1573		
Qy	1320	GATTCATCAGGGTTATTTGACTGTAATGAACTCTTTTTTGTGTTGTGTCAGCTCATAAA	1379		
Db	1574	GATTCATCAGGGTTATTTGACTGTAATGAACTCTTTTTTGTGTTGTGTCAGCTCATAAA	1633		
Qy	1380	AAATGGCTG 1388			
Db	1634	GGATCCCGG 1642			
RESULT 5	AX392814	11127 bp	DNA	linear	PAT 23-MAR-2002
LOCUS	AX392814	Sequence 17 from Patent WO0215701.			
DEFINITION	AX392814				
ACCESSION	AX392814				
VERSION	AX392814.1	GI:19700916			
KEYWORDS		synthetic construct			
SOURCE		synthetic construct			
ORGANISM		other sequences; artificial sequences.			
REFERENCE	1				

AUTHORS	Carozzi,N.B., Rabe,S.M., Miles,P.J., Warren,G.W. and de Haan,P.T.										
TITLE	Novel insecticidal toxins derived from Bacillus thuringiensis insecticidal crystal proteins										
JOURNAL	Patent: WO 0215701-A 17 28-FEB-2002; Syngenta Participations AG (CH)										
FEATURES	Location/Qualifiers										
source	1..11127	/organism="synthetic construct"									
		/mol_type="unassigned DNA"									
		/db_xref="taxon:32630"									
		/note="pZU578"									
misc_feature	56..1475	/note="Actin 2 promoter U41998"									
misc_feature		1485..3491	/note="synthetic nucleotide sequence encoding the toxin portion of H04 plus the first 40 amino acids of the CryIab tail"								
misc_feature		3859..5030	/note="SWAS promoter"								
misc_feature		5052..6271	/note="PMI"								
ORIGIN											
Query Match	90.5%; Score 1306.2; DB 6; Length 11127;										
Best Local Similarity	98.8%; Pred. No. 8.3e-272;										
Matches 1368; Conservative	0; Mismatches 10; Indels 7; Gaps 5;										
QY	1	ATTATGATCTCAAAATCATTTGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG	60								
Db	93	ATTATGATCTCAAAATCATTTGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG	152								
QY	61	AAAGTTTTGACGAATATGNNACGACAAAATGGCTACATCGATGTAATTTGGTATCTCAAC	120								
Db	153	AAAGTTTTGACGAATATGGCAGCAAAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC	212								
QY	121	TCAACATTATCTTATACCAACATTAGTTAG-CAAAAATTTAAACAACATA-TTTTATGT	178								
Db	213	TCAACATTATCTTATACCAACATTAGTTAGCAAAAATTTAAACAACATA-TTTTATGT	272								
QY	179	ATGCAAGAGTCAGCATATGTATAATTTGATTTCAGAATCGTTTTCACGAGTTCGGATGTAGT	238								
Db	273	ATGCAAGAGTCAGCATATGTATAATTTGATTTCAGAATCGTTTTCACGAGTTCGGATGTAGT	332								
QY	239	AGTAGCCATTTATTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGTAT	297								
Db	333	AGTAGCCATTTATTAATGTACATATCTAATCGTGAATAGTGAAATATGATGAACATTTGTAT	392								
QY	298	CTTATTGTATAAATATCCATAAACACATCATGAAAGACACTTTCTTTCACGGTCTGAATT	357								
Db	393	CTTATTGTATAAATATCCATAAACACATCATGAAAGACACTTTCTTTCACGGTCTGAATT	452								
QY	358	AATTATGATACAAATCTTAATAGAAAACGAATTAATTAATTAAGTTGAATTTGATGAATCTAA	417								
Db	453	AATTATGATACAAATCTTAATAGAAAACGAATTAATTAATTAAGTTGAATTTGATGAATCTAA	512								
QY	418	TTGAACAAGCCCAACACGACGAGGACTAACCGTTGCGCTGGATGACTTCGGTTTAAAGTTAAC	477								
Db	513	TTGAACAAGCCCAACACGACGAGGACTAACCGTTGCGCTGGATGACTTCGGTTTAAAGTTAAC	572								
QY	478	CACATAAAAAACGGAGCTGTCTATGTAACAACGCGGATCCGAGGTCACAGTCATGAAGCC	537								
Db	573	CACATAAAAAACGGAGCTGTCTATGTAACAACGCGGATCCGAGGTCACAGTCATGAAGCC	632								
QY	538	ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTAAGTTTAAAAAATTAGTTAA	597								
Db	633	ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTAAGTTTAAAAAATTAGTTAA	692								
QY	598	CACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACTGTTATCTTTACCTGTGGTCGAAATGA	656								
Db	693	CACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACTGTTATCTTTACCTGTGGTCGAAATGA	752								
QY	657	TTCTGTCTGTCTGATTTTAAATTTTATTTTGAAGCCGCAAAATAAAGTTTGAAGATA	716								

Db	753	TTCTGTCGTCGATTTTAATTAATTTTGTGAAAGCCGAAATAAAGTTGTGAAGATA	812
Qy	717	AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAATGTCTGTTGCTCTCC	776
Db	813	AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAATGTCTGTTGCTCTCC	872
Qy	777	TCACTTTTCATCAGCGCTTTTGAATCTCCGCGACTTTGACAGAGAACAAGGAAGA	836
Db	873	TCACTTTTCATCAGCGCTTTTGAATCTCCGCGACTTTGACAGAGAACAAGGAAGA	932
Qy	837	CTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCTCTCCGTTTGAATCTCTCTCAA	896
Db	933	CTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCTCTCCGTTTGAATCTCTCTCAA	992
Qy	897	TCTCATCTCTCTCTCTCCGCTCTCTCTCTCCAGGTAATAGGAATCTCTGGAATCTACTTT	956
Db	993	TCTCA---TCTTCTCTCCGCTCTCTCTCTCCAGGTAATAGGAATCTCTGGAATCTACTTT	1049
Qy	957	ATTTCGTGATCTCGATCTTGTCTCTCTCAATTTCTTGAGATCTGGAATCTGTTAAATTT	1016
Db	1050	ATTTCGTGATCTCGATCTTGTCTCTCTCAATTTCTTGAGATCTGGAATCTGTTAAATTT	1109
Qy	1017	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGAATCTGAATCTGACCGAT	1076
Db	1110	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGAATCTGAATCTGACCGAT	1169
Qy	1077	CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCATGTTTC	1136
Db	1170	CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCATGTTTC	1229
Qy	1137	ATGTTACCTCGGAATGATTTGATATGTAATGAAATCTGAATCTGAATCTGAATCTGATTT	1196
Db	1230	ATGTTACCTCGGAATGATTTGATATGTAATGAAATCTGAATCTGAATCTGATTT	1289
Qy	1197	GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAGTTAGATGAATTT	1256
Db	1290	GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAGTTAGATGAATTT	1349
Qy	1257	GTGATAGATTTCTCGAAACCTTAGAATTTGATGTTCTGATGTTGAACAGAAAGCTATT	1316
Db	1350	GTGATAGATTTCTCGAAACCTTAGAATTTGATGTTCTGATGTTGAACAGAAAGCTATT	1409
Qy	1317	TCTGATTCATCAGGGTTTATTTGACTGTATTTGAATCTGATTTGATTTGTTGTTTCAGCTCAT	1376
Db	1410	TCTGATTCATCAGGGTTTATTTGACTGTATTTGAATCTGATTTGATTTGTTGTTTCAGCTCAT	1469
Qy	1377	AAAAA 1381	
Db	1470	AAAAA 1474	
RESULT 6			
CQ774496			
LOCUS	CQ774496 14184 bp DNA linear PAT 06-MAR-2004		
DEFINITION	Sequence 58 from Patent WO2004013333.		
ACCESSION	CQ774496		
VERSION	CQ774496.1 GI:45237730		
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Kock, M., Frank, M. and Badur, R.		
TITLE	Novel selection method		
JOURNAL	Patent: WO 2004013333-A 58 12-FEB-2004;		
FEATURES	BASF Plant Science GmbH (DE)		
	Location/Qualifiers		
	1..14184		
	/organism="synthetic construct"		
source	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Beschreibung der kuenstlichen Sequenz: Expression		
	vector pSUN1-coda-RNAi-Act.-2-Act.-R-ocsr"		

ORIGIN

Query Match		90.1%;	Score 1300.8;	DB 6;	Length 14184;
Best Local Similarity		98.1%;	Pred. No. 1.2e-270;		
Matches 1379;		Conservative 0;	Mismatches 19;	Indels 8;	Gaps 6;
Qy	1	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG	60		
Db	9980	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG	10039		
Qy	61	AAAGTTTTTCCAGCAATATGNNACGACAAAATGGTACACCTCGATGTAATTTGGTATCTCAAC	120		
Db	10040	AAAGTTTTTCCAGCAATATGNNACGACAAAATGGTACACCTCGATGTAATTTGGTATCTCAAC	10099		
Qy	121	TCAACATTTATCTTATACCAAAATTTAGTTAG-CAAAAATTTAAACAACATA-TTTTATATGT	178		
Db	10100	TCAACATTTATCTTATACCAAAATTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT	10159		
Qy	179	ATGCAAGAGTCAAGCATATGTAATTTGATTGCAATCGTTTTCAGAGTTCGATGTAGT	238		
Db	10160	ATGCAAGAGTCAAGCATATGTAATTTGATTGCAATCGTTTTCAGAGTTCGATGTAGT	10219		
Qy	239	AGTAGCCATTATTTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGAT	297		
Db	10220	AGTAGCCATTATTTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGAT	10279		
Qy	298	CTTATTGTATAAATATCCATAAACACATCATGAAAGACACTTTTTCACAGGCTCTGAAATT	357		
Db	10280	CTTATTGTATAAATATCCATAAACACATCATGAAAGACACTTTTTCACAGGCTCTGAAATT	10339		
Qy	358	AATTATGATACAAATTTCTAATAGAAAACGAATTTAAATTTAGCTTGAATTTGTAATCTAA	417		
Db	10340	AATTATGATACAAATTTCTAATAGAAAACGAATTTAAATTTAGCTTGAATTTGTAATCTAA	10399		
Qy	418	TTGAACAAGCCCAACCAAGCAGGAGTAACTACGTTGCTGGAATTTGACTTCGGTTAAAGTTAAC	477		
Db	10400	TTGAACAAGCCCAACCAAGCAGGAGTAACTACGTTGCTGGAATTTGACTTCGGTTAAAGTTAAC	10459		
Qy	478	CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTCACAGTCAATGAAGCC	537		
Db	10460	CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTCACAGTCAATGAAGCC	10519		
Qy	538	ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	597		
Db	10520	ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	10579		
Qy	598	CACGAGGGGAAA-GCTGTCTGACAGCCAGGTCACGTTGCTGGAATTTGACTTCGGTTAAAGTTAAC	656		
Db	10580	CACGAGGGGAAAAGGCTGTCTGACAGCCAGGTCACGTTTATCTTTTACCTGTGTCGAAATGA	10639		
Qy	657	TTCTGTCTGTCTCGATTTTAAATTTTAAAGGCGGAAATAAAGTTGTAAAGATA	716		
Db	10640	TTCTGTCTGTCTCGATTTTAAATTTTAAAGGCGGAAATAAAGTTGTAAAGATA	10699		
Qy	717	AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAATTTGATTTGCTGTTGCTCTCC	776		
Db	10700	AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAATTTGATTTGCTGTTGCTCTCC	10759		
Qy	777	TCACCTTCATCAGCGCTTTTGAATCTCCGCGGACTTTGACAGAGAGAAACAAGAGAGA	836		
Db	10760	TCACCTTCATCAGCGCTTTTGAATCTCCGCGGACTTTGACAGAGAGAAACAAGAGAGA	10819		
Qy	837	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCTCTCCGTTTGAATCTCTCTCAA	896		
Db	10820	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCTCTCCGTTTGAATCTCTCTCAA	10879		
Qy	897	TCTCATCTCTCTCTCTCCGCTCTCTCTCTCCAGGTAATAGGAATCTCTGGAATCTACTTT	956		
Db	10880	TCTCA---TCTTCTCTCCGCTCTCTCTCTCCAGGTAATAGGAATCTCTGGAATCTACTTT	10936		
Qy	957	ATTTCGTGATCTCGATCTTGTCTCTCAATTTCTTGAGATCTGGAATCTGTTAAATTT	1016		
Db	10937	ATTTCGTGATCTCGATCTTGTCTCTCAATTTCTTGAGATCTGGAATCTGTTAAATTT	10996		

Qy	1017	GGATCTGTGAACTCCACATAAATCTTTTGGTTTACTAGAACTGATCTAAAGTTGACCGAT	1076
Db	10997	GGATCTGTGAACTCCACATAAATCTTTTGGTTTACTAGAACTGATCTAAAGTTGACCGAT	11056
Qy	1077	CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCATGTTTC	1136
Db	11057	CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCATGTTTC	11116
Qy	1137	ATGTTTACCTGGGAATGATTTGTATATGTAATTCGAAATCTGAACTGTTGAAAGTTAGATT	1196
Db	11117	ATGTTTACCTGGGAATGATTTGTATATGTAATTCGAAATCTGAACTGTTGAAAGTTAGATT	11176
Qy	1197	GAATCTGAACTGTCGAATGTTTAGATTGAATCTGAACTGTTTAA-GTTAGATGAAGTT	1255
Db	11177	GAATCTGAACTGTCGAATGTTTAGATTGAATCTGAACTGTTTAAAGGTTAGATGAAGTT	11236
Qy	1256	TGTGTATAGATCTTCCGAACTTAGGATTTGTAGTGTGCTGATGAACTGAACTAGCTAT	1315
Db	11237	TGTGTATAGATCTTCCGAACTTAGGATTTGTAGTGTGCTGATGAACTGAACTAGCTAT	11296
Qy	1316	TTCTGATTCGAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTGTTTGCAGCTCA	1375
Db	11297	TTCTGATTCGAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTGTTTGCAGCTCA	11356
Qy	1376	TAAAAAATGGCTGAGGCTCAGCAT 1401	
Db	11357	TAAAAAATGGCTGAGGCTCAGGCAT 11382	
RESULT 7			
LOCUS	AX463620	3450 bp	DNA linear PAT 15-JUL-2002
DEFINITION	Sequence 17 from Patent WO0248335.		
ACCESSION	AX463620		
VERSION	AX463620.1	GI:21886380	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Meagher,R.B. and Li,Y.		
TITLE	Metal resistant plants and phytoremediation of environmental contamination		
JOURNAL	Patent: WO 0248335-A 17 20-JUN-2002;		
FEATURES	UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)		
source	Location/Qualifiers		
	1..3450		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Plasmid pECSACT2"		
ORIGIN			
Query Match	89.9%;	Score 1298.4;	DB 6; Length 3450;
Best Local Similarity	98.1%;	Pred. No. 4.6e-270;	
Matches 1376;	Conservative	0; Mismatches 18; Indels 8; Gaps 6;	
Qy	1	ATTATGATCTCAAAATCAATTGATACATCTCATCTAGATCTAGGTTATCATATGTAAG	60
Db	9	ATGCTGATCTCAAAATCAATTGATACATCTCATCTAGATCTAGGTTATCATATGTAAG	68
Qy	61	AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC	120
Db	69	AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC	128
Qy	121	TCAACATTATCTTATACCAAACTTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATGTT	178
Db	129	TCAACATTATCTTATACCAAACTTAGTTAGCAAAAAATTTAAACAACCTA-TTTTATGTT	188
Qy	179	ATGCAAGAGTTCAGCATATGTATTAATTTGATTCAGAATCGTTTTCAGAGTTCGATGTAGT	238
Db	189	ATGCAAGAGTTCAGCATATGTATTAATTTGATTCAGAATCGTTTTCAGAGTTCGATGTAGT	248

QY	239	AGTAGCCATTATTTAATGTACATCTAATAATCGTGAATAGTG-ATATGATGAAACATTTGTAT	297
Db	249	AGTAGCCATTATTTAATGTACATCTAATAATCGTGAATAGTGATATGATGAACATTTGTAT	308
QY	298	CTTATTGTATAAATATCCATAAACAATCATGAAAGACACATTTTCTTTCCAGGCTCGAATT	357
Db	309	CTTATTGTATAAATATCCATAAACAATCATGAAAGACACATTTTCTTTCCAGGCTCGAATT	368
QY	358	AAATTATGATACAAATTTCTAATAGAAAAAGAAATTAATAATTAGTTGAATTTGATGAAATCTTAA	417
Db	369	AAATTATGATACAAATTTCTAATAGAAAAAGAAATTAATAATTAGTTGAATTTGATGAAATCTTAA	428
QY	418	TTGAACAAGCCCAACACGACGAGGACTAAACGTTGCTCGATTGACTTCGGTTTAAAGTTAAC	477
Db	429	TTGAACAAGCCCAACACGACGAGGACTAAACGTTGCTCGATTGACTTCGGTTTAAAGTTAAC	488
QY	478	CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTCACAGTCAATGAAGCC	537
Db	489	CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTCACAGTCAATGAAGCC	548
QY	538	ATCAAAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	597
Db	549	ATCAAAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	608
QY	598	CACGAGGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTTATCTTTTACCTGTGGTCGAAATGA	656
Db	609	CACGAGGGAAAAAAGGCTGTCTGACAGCCAGGTCACGTTTATCTTTTACCTGTGGTCGAAATGA	668
QY	657	TTCTGTCTGTCTGATTTTAAATTTATTTTGTAAAGCCGAAATAAAGTTGTGAAGATGA	716
Db	669	TTCTGTCTGTCTGATTTTAAATTTATTTTGTAAAGCCGAAATAAAGTTGTGAAGATGA	728
QY	717	AACCCGCTATATAAATTTCAATATTTTCTCCCGCTTTGAAATTTGTCCTGTTGCTCTCC	776
Db	729	AACCCGCTATATAAATTTCAATATTTTCTCCCGCTTTGAAATTTGTCCTGTTGCTCTCC	788
QY	777	TCACTTTTCATCAGCGTTTGTGAATCTCCGGGACCTTGACAGAGAAAGAACAGAGAAAGA	836
Db	789	TCACTTTTCATCAGCGTTTGTGAATCTCCGGGACCTTGACAGAGAAAGAACAGAGAAAGA	848
QY	837	CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATCTTCTCTCAA	896
Db	849	CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATCTTCTCTCAA	908
QY	897	TCTCATCTTCTTCTTCCGCTCTTTCTTTCCAAAGGTAATAGGAACTTTCTGGAATCTACTTT	956
Db	909	TCTCA---TCTTCTTCCGCTCTTTCTTTCCAAAGGTAATAGGAACTTTCTGGAATCTACTTT	965
QY	957	ATTTGCTGATCTCGATCTTGTGTTTCTCAAATTTTCTTGGATCTCGAAATTCGTTTAAATTT	1016
Db	966	ATTTGCTGATCTCGATCTTGTGTTTCTCAAATTTTCTTGGATCTCGAAATTCGTTTAAATTT	1025
QY	1017	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAACTCGATCTAAGTTGACCGAT	1076
Db	1026	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAACTCGATCTAAGTTGACCGAT	1085
QY	1077	CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCATGTTTC	1136
Db	1086	CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCATGTTTC	1145
QY	1137	ATGTTTACCTGGGAAATGATTTGTATATGTAATTCGAAATCTCAACTGTTTGAAGTTAGATT	1196
Db	1146	ATGTTTACCTGGGAAATGATTTGTATATGTAATTCGAAATCTCAACTGTTTGAAGTTAGATT	1205
QY	1197	GAATCTGAACACTGTCAATTTGATGATGAATCTGAACACTGTTTAA-GTTAGATGAAGTT	1255
Db	1206	GAATCTGAACACTGTCAATTTGATGATGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT	1265
QY	1256	TGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGTTCGATGAACTAGCTAT	1315
Db	1266	TGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGTTCGATGAACTAGCTAT	1325

QY 1316 TTCTGATTCAATCAGGGTTTATTGACTGATGTAAGACTCTTTTGTGTGTTTGCAGCTCA 1375
 Db 1326 TTCTGATTCAATCAGGGTTTATTGACTGATGTAAGACTCTTTTGTGTGTTTGCAGCTCA 1385
 QY 1376 TAAAAAATGGCTGAGGCTGACG 1397
 Db 1386 TAAACNUGGAATCCCGAGC 1407

RESULT 8
 LOCUS CQ814616 7474 bp DNA linear PAT 24-MAY-2004
 DEFINITION Sequence 7 from Patent WO2004039986.
 ACCESSION CQ814616
 VERSION CQ814616.1 GI:47603799
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.

REFERENCE
 1 Ellis,D.M., Negrotto,D.V., Shi,L., Shotkoski,F.A. and Thomas,C.R.
 Cot102 insecticidal cotton
 Patent: WO 2004039986-A 7 13-MAY-2004;
 Syngenta Participations AG (CH)
 Location/Qualifiers
 i. .7474
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Cot102 nucleotide motif"

ORIGIN

Query Match 89.9%; Score 1298.4; DB 6; Length 7474;
 Best Local Similarity 98.8%; Pred. No. 4.2e-270;
 Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAATATACATGATATACATATCTCATCTAGATCTAGGTATCATATGTAAG 60
 Db 3228 ATTATGATCTCAATATACATGATATACATATCTCATCTAGATCTAGGTATCATATGTAAG 3287
 QY 61 AAAGTTTGGACCAATATGNNACGACAAATGCTACACTCGATGTAATGGTATCTCAAC 120
 Db 3288 AAAGTTTGGACCAATATGNNACGACAAATGCTACACTCGATGTAATGGTATCTCAAC 3347
 QY 121 TCAACATTATATCTTACCAACATTTAGTTAG-CAAAATTTAAACAACCTA-TTTTATATGT 178
 Db 3348 TCAACATTATATCTTACCAACATTTAGTTAGCAAAATTTAAACAACCTATTTTATGT 3407
 QY 179 ATGCAAGAGTCAGCATATGTAATTAATGATTCAGAAATCGTTTGGACGAGTTCGATGTAGT 238
 Db 3408 ATGCAAGAGTCAGCATATGTAATTAATGATTCAGAAATCGTTTGGACGAGTTCGATGTAGT 3467
 QY 239 AGTAGCCATTATTAATGTACATCTATCTGATAGT-ATATGATCAACATGTAT 297
 Db 3468 AGTAGCCATTATTAATGTACATCTATCTGATAGTATCGTGAATATGATGAACATGTAT 3527
 QY 298 CTATTGTATATAATATCCATAACACATCATCAAAAGACACTTTCTTTGAGGCTCTGAATT 357
 Db 3528 CTATTGTATATAATATCCATAACACATCATCAAAAGACACTTTCTTTGAGGCTCTGAATT 3587
 QY 358 AATTATGATACAAATCTTAATAGAAAACGAATTAATAGTTGAATGATGAAATCTAA 417
 Db 3588 AATTATGATACAAATCTTAATAGAAAACGAATTAATAGTTGAATGATGAAATCTAA 3647
 QY 418 TTGAACAAGCCAAACCAAGCAGGAGTAACGTTGCTGATGATGACTCGTTTAAAGTTAAC 477
 Db 3648 TTGAACAAGCCAAACCAAGCAGGAGTAACGTTGCTGATGATGACTCGTTTAAAGTTAAC 3707
 QY 478 CACTAAAAAAGCGAGCTCTCATGTAAACACGCGGATCGAGCAGGTCAACAGTCATGAAGCC 537
 Db 3708 CACTAAAAAAGCGAGCTCTCATGTAAACACGCGGATCGAGCAGGTCAACAGTCATGAAGCC 3767
 QY 538 ATCAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAA 597

3768 ATCAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAA 3827
 598 CACGAGGAGAAA-GCTGTCTGACAGCCAGGTACAGTATCTTTTACCTGTGGTCCGAATCA 656
 3828 CACGAGGAGAAAAGGCTGTGACAGCCAGGTACAGTATCTTTTACCTGTGGTCCGAATCA 3887
 657 TTCGTGCTGTGCGATTTTAAATTTATTTTGAAGCCGAAATAAAGTTGTAAGAGATA 716
 3888 TTCGTGCTGTGCGATTTTAAATTTATTTTGAAGCCGAAATAAAGTTGTAAGAGATA 3947
 717 AACCCGCTATATAAATTTATATAATTTTCTCCCGCTTTGAAATTTGTCGTGTCCTCC 776
 3948 AACCCGCTATATAAATTTATATAATTTTCTCCCGCTTTGAAATTTGTCGTGTCCTCC 4007
 777 TCACCTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAAACAAGGAAGA 836
 4008 TCACCTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAAACAAGGAAGA 4067
 837 CTAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAA 896
 4068 CTAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAA 4127
 897 TCTCATCTTCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTCTGGAATCTACTTT 956
 4128 TCTCA---TCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTCTGGAATCTACTTT 4184
 957 ATTGCTGAGATCTGATCTTGTGTTTCTCAATTTCTGAGATCTGGAATTCGTTTAAATTT 1016
 4185 ATTGCTGAGATCTGATCTTGTGTTTCTCAATTTCTGAGATCTGGAATTCGTTTAAATTT 4244
 1017 GGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGAT 1076
 4245 GGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGAT 4304
 1077 CAGTTAGTCTGATATAGCTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTTTC 1136
 4305 CAGTTAGTCTGATATAGCTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTTTC 4364
 1137 ATGTTACCTCGGAATGATTTGTATGTGAATCTGAAATCTGAACTGTTTGAAGTTAGATT 1196
 4365 ATGTTACCTCGGAATGATTTGTATGTGAATCTGAAATCTGAACTGTTTGAAGTTAGATT 4424
 1197 GAATCTGAACACTGTCATGTTAGATTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255
 4425 GAATCTGAACACTGTCATGTTAGATTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT 4484
 1256 TGTGTATAGATCTTCGAAACCTTAGGATTTGTAGTGTGCTACGTTGAACAGAAAGCTAT 1315
 4485 TGTGTATAGATCTTCGAAACCTTAGGATTTGTAGTGTGCTACGTTGAACAGAAAGCTAT 4544
 1316 TTCTGATTCATCAGGGTTTATTTGACTGATTTGAACCTTTTGTGTGTTTGCAGCTCA 1375
 4545 TTCTGATTCATCAGGGTTTATTTGACTGATTTGAACCTTTTGTGTGTTTGCAGCTCA 4604
 1376 TAAAAA 1381
 4605 TAAAAA 4610

RESULT 9
 LOCUS CQ814630 9356 bp DNA linear PAT 24-MAY-2004
 DEFINITION Sequence 21 from Patent WO2004039986.
 ACCESSION CQ814630
 VERSION CQ814630.1 GI:47603812
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.

REFERENCE
 1 Ellis,D.M., Negrotto,D.V., Shi,L., Shotkoski,F.A. and Thomas,C.R.
 Cot102 insecticidal cotton

JOURNAL Patent: WO 2004039986-A 21 13-MAY-2004;
Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
source 1..9356
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="COT102 nucleotide motif"

ORIGIN

Query Match 89.9%; Score 1298.4; DB 6; Length 9356;
Best Local Similarity 98.8%; Pred. No. 4.1e-270;
Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 60
DB |||||
QY 4614 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 4673
DB |||||
QY 61 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
DB |||||
QY 4674 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 4733
DB |||||
QY 121 TCAACATTATATCTATACAAACATTAGTTAG-CAAAATTTAAACAACTA-TTTTATGT 178
DB |||||
QY 4734 TCAACATTATATCTATACAAACATTAGTTAGTACAAAATTTAAACAACTATTTTATGT 4793
DB |||||
QY 179 ATGCAAGAGTCAGCATATGTAATTTGATTTCAGAAATCGTTTTCAGAGTTCCGATGTAGT 238
DB |||||
QY 4794 ATGCAAGAGTCAGCATATGTAATTTGATTTCAGAAATCGTTTTCAGAGTTCCGATGTAGT 4853
DB |||||
QY 239 AGTAGGCATTTTAAATGTACATCTAATTCGTGAATAGTG-ATATGATGAAACATTTGTAT 297
DB |||||
QY 4854 AGTAGGCATTTTAAATGTACATCTAATTCGTGAATAGTGATGATGATGATGATGAT 4913
DB |||||
QY 298 CTTATTGTATAAATPCCAATAACACATCATGAAAGACACTTTCTTTCAGGTCGTGAAT 357
DB |||||
QY 4914 CTTATTGTATAAATPCCAATAACACATCATGAAAGACACTTTCTTTCAGGTCGTGAAT 4973
DB |||||
QY 358 AATTATGATACATCTATATGAAACGAAATTAATTTAGTTGAAATTTGATGAAATCTAA 417
DB |||||
QY 4974 AATTATGATACATCTATATGAAACGAAATTAATTTAGTTGAAATTTGATGAAATCTAA 5033
DB |||||
QY 418 TTGAAACAGCCCAACACGAGGAGCTAAGCTTTCCTGATGATGATGATGATGATGATGAT 477
DB |||||
QY 5034 TTGAAACAGCCCAACACGAGGAGCTAAGCTTTCCTGATGATGATGATGATGATGATGAT 5093
DB |||||
QY 478 CACTAAAAAAGCGAGCTGTCTGTAAACACGCGGATCGAGCAGGTTCACAGTCAATGAAGCC 537
DB |||||
QY 5094 CACTAAAAAAGCGAGCTGTCTGTAAACACGCGGATCGAGCAGGTTCACAGTCAATGAAGCC 5153
DB |||||
QY 538 ATCAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 597
DB |||||
QY 5154 ATCAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 5213
DB |||||
QY 598 CACGAGGGAATAA-GCTGTCTGACAGCCAGGTCACTGTTTACCTGTGGTTCGAAATCA 656
DB |||||
QY 5214 CACGAGGGAATAAAGGCTGTCTGACAGCCAGGTCACTGTTTACCTGTGGTTCGAAATCA 5273
DB |||||
QY 657 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGGCCGAAATAAAGTTGTAAGAGATA 716
DB |||||
QY 5274 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGGCCGAAATAAAGTTGTAAGAGATA 5333
DB |||||
QY 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGCTTCGTTGCTCTCC 776
DB |||||
QY 5334 AACCCGCTATATAAATTCATATATTTTCTCTCCCGCTTTGAAATTTGCTTCGTTGCTCTCC 5393
DB |||||
QY 777 TCACCTTTCATCAGCCGCTTTTGAATCTCCGCGCATTTGACAGAGAAACAAAGGAAGAAGA 836
DB |||||
QY 5394 TCACCTTTCATCAGCCGCTTTTGAATCTCCGCGCATTTGACAGAGAAACAAAGGAAGAAGA 5453
DB |||||
QY 837 CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCAATCTCCGTTTGAATCTTCTCTCAA 896
DB |||||
QY 5454 CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCAATCTCCGTTTGAATCTTCTCTCAA 5513
DB |||||

QY 897 TCTCATCTTCTTCTTCCGCTCTTTCTTTCCAAGGTAATAGGAACCTTTCTGATCTACTTTT 956
DB |||||
QY 5514 TCTCA--TCTTCTTCCGCTCTTTCTTTCCAAGGTAATAGGAACCTTTCTGATCTACTTT 5570
DB |||||
QY 957 ATTTCCTGGAATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGGAATCTGTTAAATTT 1016
DB |||||
QY 5571 ATTTCCTGGAATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGGAATCTGTTAAATTT 5630
DB |||||
QY 1017 GGATCTGTGAACCTCCACTTAATCTTTTGGTTTACTAGAACGATCTAAGTTGACCGAT 1076
DB |||||
QY 5631 GGATCTGTGAACCTCCACTTAATCTTTTGGTTTACTAGAACGATCTAAGTTGACCGAT 5690
DB |||||
QY 1077 CAGTTAGCTCGAATATAGCTACCAAAATTTGGCTTTCACCTTGATGGAGAGATCCCATGTTTC 1136
DB |||||
QY 5691 CAGTTAGCTCGAATATAGCTACCAAAATTTGGCTTTCACCTTGATGGAGAGATCCCATGTTTC 5750
DB |||||
QY 1137 ATGTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACTGTTGAAAGTTAGATT 1196
DB |||||
QY 5751 ATGTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACTGTTGAAAGTTAGATT 5810
DB |||||
QY 1197 GAATCTGAACACTGTCAATGTTAGATTGAACTGAACTGTTTAA-GTTAGATGAAAGTT 1255
DB |||||
QY 5811 GAATCTGAACACTGTCAATGTTAGATTGAACTGAACTGTTTAAAGTTAGATGAAAGTT 5870
DB |||||
QY 1256 TGTGTATAGATTCTTTCGAAACCTTAGGATTTTGTAGTTCGTACGTTGAACAGAAAGCTAT 1315
DB |||||
QY 5871 TGTGTATAGATTCTTTCGAAACCTTAGGATTTTGTAGTTCGTACGTTGAACAGAAAGCTAT 5930
DB |||||
QY 1316 TTCTGATTAATTCAGGTTTATTTGACTGTATGAACTCTTTTGTGTTGTTGAGCTCA 1375
DB |||||
QY 5931 TTCTGATTAATTCAGGTTTATTTGACTGTATGAACTCTTTTGTGTTGTTGAGCTCA 5990
DB |||||
QY 1376 TAAAAA 1381
DB |||||
QY 5991 TAAAAA 5996
DB |||||

RESULT 10
AX463619
LOCUS AX463619
DEFINITION Sequence 16 from Patent WO0248335.
ACCESSION AX463619
VERSION AX463619.1 GI:21886379
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Meagher,R.B. and Li,Y.
TITLE Metal resistant plants and phytoremediation of environmental
JOURNAL contamination
Patent: WO 0248335-A 16 20-JUN-2002;
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)
FEATURES
source
1..3408
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pAtPCSACT2"

ORIGIN
Query Match 89.8%; Score 1297.4; DB 6; Length 3408;
Best Local Similarity 98.5%; Pred. No. 7.6e-270;
Matches 1372; Conservative 0; Mismatches 13; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 60
DB |||||
QY 9 ATGCTGATCTCAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 68
DB |||||
QY 61 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
DB |||||
QY 69 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 128
DB |||||

121 TCAACATTATCTTATACCAACATTTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATATGT 178
129 TCAACATTATCTTATACCAACATTTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATATGT 188
179 ATGCAAGAGTCAGCATATGTATTAATTTGATTTCAGAACTCGTTTTCGAGAGTTCCGAGTGTAGT 238
189 ATGCAAGAGTCAGCATATGTATTAATTTGATTTCAGAACTCGTTTTCGAGAGTTCCGAGTGTAGT 248
239 AGTAGCCATTATTTAATGTACATCTAATCGTGAATAGTG-ATATGATGAACAATTTGTAT 297
249 AGTAGCCATTATTTAATGTACATCTAATCGTGAATAGTG-ATATGATGAACAATTTGTAT 308
298 CTTATTTGTATTAATTTCCATAACACATCATGAAAGACACTTTTTCAGGAGTCTGAATTT 357
309 CTTATTTGTATTAATTTCCATAACACATCATGAAAGACACTTTTTCAGGAGTCTGAATTT 368
358 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAATCTTAA 417
369 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAATCTTAA 428
418 TTGAACAAGCCCAACACGACGAGGACTAAGCTTTCCTGGAATGACTTCGGTTTAAAGTTTAA 477
429 TTGAACAAGCCCAACACGACGAGGACTAAGCTTTCCTGGAATGACTTCGGTTTAAAGTTTAA 488
478 CACTAAAAAAGCGAGCTGTCTATGTTAAACACGCGGATCGAGAGGTCACAGTCATGAAGCC 537
489 CACTAAAAAAGCGAGCTGTCTATGTTAAACACGCGGATCGAGAGGTCACAGTCATGAAGCC 548
538 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAA 597
549 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAA 608
598 CACGAGGGAAAA-GCTGTCTGACAGCGAGGTCACAGTTATCTTTTACCTGTGGTCGAAATGA 656
609 CACGAGGGAAAAAGGCTGTCTGACAGCGAGGTCACAGTTATCTTTTACCTGTGGTCGAAATGA 668
657 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGGCGGAAAAATAAAGTTGTAAGAGATA 716
669 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGGCGGAAAAATAAAGTTGTAAGAGATA 728
717 AACCCGCTATATAATTTCTATATTTTCTCCCGCTTTGATTTGTCCTGTTGCTCTCC 776
729 AACCCGCTATATAATTTCTATATTTTCTCCCGCTTTGATTTGTCCTGTTGCTCTCC 788
777 TCACCTTTTCATCAGCCGTTTGTGAATCTCCGCGACTTTCAGAGAGAGAAACAAGAGAGAGA 836
789 TCACCTTTTCATCAGCCGTTTGTGAATCTCCGCGACTTTCAGAGAGAGAAACAAGAGAGAGA 848
837 CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTTGAATCTTCTCTCAA 896
849 CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTTGAATCTTCTCTCAA 908
897 TCCTATCTTCTTCTCCGCTCTTTCTTTCCAGGTAATAGGAACTTTCTGGATCTACTTTT 956
909 TCTCA--TCTTCTTCCGCTCTTTCTTTCCAGGTAATAGGAACTTTCTGGATCTACTTTT 965
957 ATTTCTGGATCTCGAATCTTGTTTTCTCAATTTCTTGTAGATCTCGAATTCGTTTAAATTT 1016
966 ATTTCTGGATCTCGAATCTTGTTTTCTCAATTTCTTGTAGATCTCGAATTCGTTTAAATTT 1025
1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAATTCGATCTAAGTTGACCGAT 1076
1026 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAATTCGATCTAAGTTGACCGAT 1085
1077 CAGTTAGCTCGAATATAGTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTTTC 1136
1086 CAGTTAGCTCGAATATAGTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTTTC 1145
1137 ATGTTACCTGGGAATGATTTGTATATGTAATTCGAATTCGAATCTGAACTGTTGAGTTAGATT 1196
1146 ATGTTACCTGGGAATGATTTGTATATGTAATTCGAATTCGAATCTGAACTGTTGAGTTAGATT 1205

1197 GAATCTGAACACCTGTCATATGTTAGATTGAATCTGAACACCTGTTTAA-GTTAGATGAAGTT 1255
1206 GAATCTGAACACCTGTCATATGTTAGATTGAATCTGAACACCTGTTTAAAGTTTAGATGAAGTT 1265
1256 TGTGTATAGATTCTTCGAAACCTTACGAGTTCGTAGTTCGTAGTTCGAGTTCGAGAGAGCTAT 1315
1266 TGTGTATAGATTCTTCGAAACCTTACGAGTTCGTAGTTCGTAGTTCGAGTTCGAGAGAGCTAT 1325
1316 TTCTGATTCAATCAGGCTTTTATTTGACTGTATTTGAACCTCTTTTGTGTGTTTGCAGCTCA 1375
1326 TTCTGATTCAATCAGGCTTTTATTTGACTGTATTTGAACCTCTTTTGTGTGTTTGCAGCTCA 1385
1376 TAAAAAATGGCTG 1388
1386 TAAACCATGGCAG 1398
RESULT 11
AX463621
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.2857
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pSACT2"
Query Match 89.8%; Score 1297; DB 6; Length 2857;
Best Local Similarity 98.6%; Pred. No. 9.5e-270;
Matches 1371; Conservative 0; Mismatches 12; Indels 8; Gaps 6;
1 ATTATGATCTCAATACATTCATATATCTCATCTAGATCTAGCTATCATTTATGTATAG 60
9 ATCTGATCTCAATACATTCATATCTCATCTAGATCTAGCTATCATTTATGTATAG 68
61 AAAAGTTTTCGCAATATGNNACGACAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
69 AAAAGTTTTCGCAATATGNNACGACAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC 128
121 TCAACATTATCTTATACCAACATTTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATATGT 178
129 TCAACATTATCTTATACCAACATTTAGTTAGCAAAAATTTAAACAACCTA-TTTTATATGT 188
179 ATGCAAGAGTCAGCATATGTATTAATTTGATTTCAGAACTCGTTTTCGAGAGTTCCGAGTGTAGT 238
189 ATGCAAGAGTCAGCATATGTATTAATTTGATTTCAGAACTCGTTTTCGAGAGTTCCGAGTGTAGT 248
239 AGTAGCCATTATTTAATGTACATCTAATCGTGAATAGTG-ATATGATGAACAACATTTGTAT 297
249 AGTAGCCATTATTTAATGTACATCTAATCGTGAATAGTG-ATATGATGAACAACATTTGTAT 308
298 CTTATTTGTATTAATTTCCATAACACATCATGAAAGACACTTTTTCAGGAGTCTGAATTT 357
309 CTTATTTGTATTAATTTCCATAACACATCATGAAAGACACTTTTTCAGGAGTCTGAATTT 368
358 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAATCTTAA 417
369 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAATCTTAA 428

Db	721	TGAATCTTCCTCAATCTCA--TCTTCTCCGCTCTTTCTTTCCAAAGGTAATAGGAACTT	777
Qy	943	TCGGATCTACTTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTGAGATCTGG	1002
Db	778	TCGGATCTACTTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTGAGATCTGG	837
Qy	1003	AATTCGTTTAAATTTGGATCTGTGAACTCCACTTAATCTTTTGGTTTACTAGAAATCGAT	1062
Db	838	AATTCGTTTAAATTTGGATCTGTGAACTCCACTTAATCTTTTGGTTTACTAGAAATCGAT	897
Qy	1063	CTAAGTTGACCGATCAGTTAGCTGATATAGCTACCAAGAAATTTGGCTTGAATCTGATGG	1122
Db	898	CTAAGTTGACCGATCAGTTAGCTGATATAGCTACCAAGAAATTTGGCTTGAATCTGATGG	957
Qy	1123	AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1182
Db	958	AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1017
Qy	1183	GTTGAAGTTAGATTTGAATCTGAACTGTCAATGTTTAGATTTGAATCTGAACTGTTTAA	1242
Db	1018	GTTGAAGTTAGATTTGAATCTGAACTGTCAATGTTTAGATTTGAATCTGAACTGTTTAA	1077
Qy	1243	GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGAATTTGATGTTGCTAGTTG	1302
Db	1078	GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTTAGGAATTTGATGTTGCTAGTTG	1137
Qy	1303	AACAGAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTATTCGAATCTTTTGTG	1362
Db	1138	AACAGAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTATTCGAATCTTTTGTG	1197
Qy	1363	TGTTTGCAGCTCATAAAAAA 1382	
Db	1198	TGTTTGCAGCTCATAAAAAA 1217	
RESULT 14			
AR236191			
LOCUS	AR236191	1219 bp	DNA linear PAT 20-DEC-2002
DEFINITION	Sequence 9 from patent US 6462258.		
ACCESSION	AR236191		
VERSION	AR236191.1 GI:27280000		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1219)		
AUTHORS	Fincher, K.L. and Wilkinson, J.Q.		
TITLE	Plant expression constructs		
JOURNAL	Patent: US 6462258-A 9 08-OCT-2002;		
FEATURES	Location/Qualifiers		
source	1..1219		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	81.4%; Score 1174.8; DB 6; Length 1219;		
Best Local Similarity	98.7%; Pred. No. 2.6e-243;		
Matches 1206; Conservative	0; Mismatches 12; Indels 4; Gaps 2;		
Qy	164	CAACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTGGAC	223
Db	1	CAACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTGGAC	60
Qy	224	GAGTTCGGATGTAGTAGTACCATTTATTTAATGTTACATCTAATCGTGAATAGTATATG	283
Db	61	GAGTTCGGATGTAGTAGTACCATTTATTTAATGTTACATCTAATCGTGAATAGTATATG	120
Qy	284	ATGAACATTTGATCTTATTTGTATAAATATCCATAAACAATCATCAAGAACACTTTTCTT	343
Db	121	ATGAACATTTGATCTTATTTGTATAAATATCCATAAACAATCATCAAGAACACTTTTCTT	180
Qy	344	TCAGGGTCTGAAATTAATATGATACAAATCTTAATAGAAAAAGAAATTAATACGTTGAAT	403

RESULT 15
AR438153

Db	181	TCACGGTCTGAATTAATTAATGATACAAATTTCTAATAGAAAAACGAAATTAATTAACGTTGAAT	240
Qy	404	TGTATGAATCTTAATTTGAACAGCCAAACACGACGAGGACTAAACGTTGCCTGGATTGACT	463
Db	241	TGTATGAATCTTAATTTGAACAGCCAAACACGACGAGGACTAAACGTTGCCTGGATTGACT	300
Qy	464	CGTTTAAAGTTAAACCACTAAAAAAACGGAGCTGTCTATGTAAACACGCGGATCGAGCAGGTC	523
Db	301	CGTTTAAAGTTAAACCACTAAAAAAACGGAGCTGTCTATGTAAACACGCGGATCGAGCAGGTC	360
Qy	524	ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATTAATTAATAGTT	583
Db	361	ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATTAATTAATAGTT	420
Qy	584	TAAAAATTAAGTTAAACACGAGGAAAAA-GCTGTCTCAGACGAGGTCAGGTTATCTTTACC	642
Db	421	TAAAAATTAAGTTAAACACGAGGAAAAAAGGCTGTCTCAGACGAGGTCAGGTTATCTTTACC	480
Qy	643	TGTGTCTCGAAATGATTCGTGTCTCGATTTTAAATTAATTTTGAAGGCCGAAAAATAA	702
Db	481	TGTGTCTCGAAATGATTCGTGTCTCGATTTTAAATTAATTTTGAAGGCCGAAAAATAA	540
Qy	703	AGTTGTAGAGATAAAACCCGCTATATAAATTCATAATATTTTCTCCCGCTTTGAAATTG	762
Db	541	AGTTGTAGAGATAAAACCCGCTATATAAATTCATAATATTTTCTCCCGCTTTGAAATTG	600
Qy	763	TCTCGTTGCTCCTCCTCAGCTTTCATCAGCGGTTTGAATCTCCGCGGACTTGACAGAGAAG	822
Db	601	TCTCGTTGCTCCTCCTCAGCTTTCATCAGCGGTTTGAATCTCCGCGGACTTGACAGAGAAG	660
Qy	823	AACAAGGAAGAGACTAAGAGAGAAAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT	882
Db	661	AACAAGGAAGAGACTAAGAGAGAAAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT	720
Qy	883	TGAATCTTCTCAATCTCATCTTCTTCTCCGCTCTTTTCCAAAGGTAATAGGAATCTT	942
Db	721	TGAATCTTCTCAATCTCA--TCTTCTTCCGCTCTTTTCCAAAGGTAATAGGAATCTT	777
Qy	943	TCTGGATCTACATTTATTTGCTGGATCTCGATCTGTTTCTCAATTTCTCAGATCTGG	1002
Db	778	TCTGGATCTACATTTATTTGCTGGATCTCGATCTGTTTCTCAATTTCTTGGAGATCTGG	837
Qy	1003	AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAATCGAT	1062
Db	838	AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAATCGAT	897
Qy	1063	CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTACGAGAAATTTGGCTTGACCTTGATGG	1122
Db	898	CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTACGAGAAATTTGGCTTGACCTTGATGG	957
Qy	1123	AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTAATCGAAATCTGAAT	1182
Db	958	AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTAATCGAAATCTGAAT	1017
Qy	1183	GTTGAAGTTAGATTTGAATCTGAACTGTCAATGTTTAGATTTGAATCTGAACACTGTTTAA	1242
Db	1018	GTTGAAGTTAGATTTGAATCTGAACTGTCAATGTTTAGATTTGAATCTGTTTAA	1077
Qy	1243	GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGAATTTGATGTTGCTAGCTTG	1302
Db	1078	GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGAATTTGATGTTGCTAGCTTG	1137
Qy	1303	ACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGATTTGAATCTTTTGTG	1362
Db	1138	ACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGATTTGAATCTTTTGTG	1197
Qy	1363	TGTTTGCAGCTCATAAAAAATG 1384	
Db	1198	TGTTTGCAGCAGACTCACCATG 1219	

LOCUS	AR438153	1219 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	Sequence 9 from patent US 6660911.				
ACCESSION	AR438153				
VERSION	AR438153.1 GI:40204655				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1219)				
AUTHORS	Fincher,K.L., Flasinaki,S. and Wilkinson,J.Q.				
TITLE	Plant expression constructs				
JOURNAL	Patent: US 6660911-A 9 09-DEC-2003;				
FEATURES	Location/Qualifiers				
source	1..1219				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN	Query Match 81.4%; Score 1174.8; DB 6; Length 1219; .				
	Best Local Similarity 98.7%; Pred. No. 2.6e-243;				
	Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;				
QY	164	CAACTATTTTATGATGCAAGAGTCAGCATATGTATATATGATTCAGAAATCGTTTGGAC	223		
Db	1	CAACTATTTTATGATGCAAGAGTCAGCATATGTATATATGATTCAGAAATCGTTTGGAC	60		
QY	224	GAGTTCCGATAGTAGTAGCCATTATTTAAATGTACATCTAATCGTGAATAGTGATATG	283		
Db	61	GAGTTCCGATAGTAGTAGCCATTATTTAAATGTACATCTAATCGTGAATAGTGATATG	120		
QY	284	ATGAACATGTATCTTATTTATTAATATCAATAACAATCATCAAGAAAGACATTTCTTT	343		
Db	121	ATGAACATGTATCTTATTTATTAATATCAATAACAATCATCAAGAAAGACATTTCTTT	180		
QY	344	TCAGGCTCTGAATTAATTAATGATACAACTCTAATAGAAAACCAATTAATTAACGTTGAAT	403		
Db	181	TCAGGCTCTGAATTAATTAATGATACAACTCTAATAGAAAACCAATTAATTAACGTTGAAT	240		
QY	404	TGTATGAATCTAAATTTGAACCAACCAACGACGAGGACTAACGTTGCTCGATTTGACT	463		
Db	241	TGTATGAATCTAAATTTGAACCAACCAACGACGAGGACTAACGTTGCTCGATTTGACT	300		
QY	464	CGGTTTAAGTTAACCACTAAAAAACGGAGCTGTCTATGTAAACACGCGGATCCAGCAGGTC	523		
Db	301	CGGTTTAAGTTAACCACTAAAAAACGGAGCTGTCTATGTAAACACGCGGATCCAGCAGGTC	360		
QY	524	ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAGGGCTGAGATGAATTAATTAGTT	583		
Db	361	ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAGGGCTGAGATGAATTAATTAGTT	420		
QY	584	TAAAAATTAGTTAAACACGAGGGAAAAA-GCTGCTGCACAGCCAGGTCACGTTATCTTTACC	642		
Db	421	TAAAAATTAGTTAAACACGAGGGAAAAAGGCTGCTGCACAGCCAGGTCACGTTATCTTTACC	480		
QY	643	TGTGTCGAAATGATTCGTGTCGTGCAATTTAAATTAATTTTTTGAAGGCCGAAAAATAA	702		
Db	481	TGTGTCGAAATGATTCGTGTCGTGCAATTTAAATTAATTTTTTGAAGGCCGAAAAATAA	540		
QY	703	AGTTGCTAGAGATTAACCCGCTATATAATTAATTAATTAATTTCTCCCGCTTTGCAATG	762		
Db	541	AGTTGCTAGAGATTAACCCGCTATATAATTAATTAATTAATTTCTCCCGCTTTGCAATG	600		
QY	763	TCTCGTTGCTCTCTCACCTTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAAG	822		
Db	601	TCTCGTTGCTCTCTCACCTTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAAG	660		
QY	823	AACAAGAAAGAACTAAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT	882		
Db	661	AACAAGAAAGAACTAAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT	720		
QY	883	TGAATCTCTCAATCTCATCTTCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAATTT	942		
Db	721	TGAATCTCTCTCAATCTCA---TCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAATTT	777		

Search completed: May 27, 2005, 22:19:52
Job time : 7829 secs

QY	943	TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTCAATTTCTTGGATCTGG	1002
Db	778	TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTCAATTTCTTGGATCTGG	837
QY	1003	AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT	1062
Db	838	AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT	897
QY	1063	CTAAGTTGACCGATCAGTTAGCTCGATATATAGCTACCCAGAAATTTGGCTTGACCTTGATCG	1122
Db	898	CTAAGTTGACCGATCAGTTAGCTCGATATATAGCTACCCAGAAATTTGGCTTGACCTTGATCG	957
QY	1123	AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1182
Db	958	AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1017
QY	1183	GTTGAAGTTAGATTGAATCTGNAACATCTCAATGTAGATTGAATCTGNAACATCTGTTAA	1242
Db	1018	GTTGAAGTTAGATTGAATCTGNAACATCTGCAATGTAGATTGAATCTGNAACATCTGTTAA	1077
QY	1243	GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGACGTTG	1302
Db	1078	GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGACGTTG	1137
QY	1303	AACAGAAAGCTATTTCTGATTCAATCAGGGTTTATTTGACTGTATTTGAATCTTTTGTG	1362
Db	1138	AACAGAAAGCTATTTCTGATTCAATCAGGGTTTATTTGACTGTATTTGAATCTTTTGTG	1197
QY	1363	TGTTTGCAGCTCATAAAAAATG	1384
Db	1198	TGTTTGCAGCTCATAAAAAATG	1219

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 12:15:42 ; Search time 848 Seconds
(without alignments)
10080.314 Million cell updates/sec

Title: US-09-868-744B-1
Perfect score: 1444
Sequence: 1 attatgatctcaaatatcatt.....gtactggaatgtaggatcc 1444

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	99.9	1444	3	AA61367 Arabidops
2	1341.8	92.9	6385	6	ABK89341 Plasmid p
3	1341.4	92.9	4526	6	ABK89340 Plasmid p
4	1306.2	90.5	11127	6	AA036967 Arabidops
5	1300.8	90.1	14184	12	ADL27876 E coli co
6	1298.4	89.9	3450	6	ABK89343 Plasmid p
7	1298.4	89.9	7474	12	ADN61599
8	1298.4	89.9	9356	12	ADN61613
9	1297.4	89.8	3408	6	ABK89342 Plasmid p
10	1297	89.8	2857	6	ABK89344
11	1174.8	81.4	1219	4	AA009791 Arabidops
12	1174	81.3	1742	4	AA009812 Chimeric
13	1114.6	77.2	12304	8	ABV75876 Luciferas
14	1104.6	76.5	1259	6	ABK52080 Modified
15	1078.8	74.7	15676	3	AA01288 Arabidops
16	1078.8	74.7	17111	3	AA01289 Arabidops
17	1078.8	74.7	17116	3	AA01290 Arabidops
18	1053	72.9	1202	6	ABK52077 Modified
19	1001.8	69.4	1342	6	ABK52081 Modified
20	950.2	65.8	1285	6	ABK52078 Modified

21	784.4	54.3	910	6	ABK52079 Modified
22	732.8	50.7	853	6	ABK52076 Modified
23	707.2	49.0	8573	12	ADO47655 Control o
24	579	40.1	1228	6	ABK53111 Transgene
25	475.8	33.0	11461	4	AA02175 Plasmid p
26	474.8	32.9	12766	4	AA02174 Plasmid p
27	163	11.3	1486	3	AA033486 Arabidops
28	116.4	8.1	1228	6	ABK53111 Transgene
29	110	7.6	573	10	ABK56906 Arabidops
30	94	6.5	94	6	ABK52089 Synthetic
31	84.4	5.8	97	6	ABK52087 Synthetic
32	79	5.5	79	6	ABK52084 Synthetic
33	77.4	5.4	79	6	ABK52092 Synthetic
34	76	5.3	2000	6	ABZ15497 Arabidops
35	75.4	5.2	1271	4	AA009792 Arabidops
36	73.8	5.1	1800	4	AA009811 Chimeric
37	65.4	4.5	67	6	ABK52085 Synthetic
38	65.4	4.5	74	6	ABK52082 Synthetic
39	63.4	4.4	77	6	ABK52083 Synthetic
40	59.2	4.1	1975	13	ADR73222 Thale cre
41	56.2	3.9	436	12	ADP93490 Cotton ex
42	56.2	3.9	2000	8	ADA71938 Rice gene
43	52.2	3.6	633	13	ACN52368 Cotton an
44	51.6	3.6	580	13	ADR63253 Cotton cd
45	51.2	3.5	1134	3	AAC45521 Arabidops

ALIGNMENTS

RESULT 1
AA61367
ID AA61367 standard; DNA; 1444 BP.
AC AA61367;
XX
XX
DT 06-NOV-2000 (first entry)
XX
DE Arabidopsis thaliana actin promoter region.
XX
KW Actin 2; ACT2; promoter; oxalate oxidase; OXOX; insecticidal; fungicidal;
KW antiviral; stability; ds.
XX
OS Arabidopsis thaliana.
XX
XX
FH Key Location/Qualifiers
FT primer_bind complement(1..21)
FT primer_bind /tag= a
FT primer_bind 1360..1379 /tag= b
FT misc_signal 1382..1384 /tag= c
FT /note= "translation start site for actin 2"
XX
XX WO200037661-A1.
XX
XX 29-JUN-2000.
XX
XX 16-DEC-1999; 99WO-GB004317.
XX
XX 21-DEC-1998; 98GB-00028201.
XX
XX (ADVA-) ADVANTA TECHNOLOGY LTD.
XX
XX Van Dun CMP, Schepers FMA, Pertijs JH;
XX WPI; 2000-442681/38.
XX
XX Producing recombinant Compositae with increased transformation stability
XX comprises linking the DNA construct to the ACT2 gene promoter.
XX
XX Claim 2; Fig 3; 22pp; English.
XX

CC The present sequence is the promoter region and part of the coding
CC sequence of the Arabidopsis gene, actin 2 (ACT2). The promoter may be
CC used in a heterologous DNA construct to drive expression of RNA. This is
CC useful for the production of recombinant plants containing genes with
CC insecticidal, fungicidal or antiviral activity. The advantage of using
CC the ACT2 gene promoter is that it increases the stability of the
CC integrated DNA. In the present invention the ACT2 promoter was used to
CC drive expression of the wheat oxalate oxidase gene
XX
SQ Sequence 1444 BP; 451 A; 241 C; 272 G; 478 T; 0 U; 2 Other;
Query Match 99.9%; Score 1442; DB 3; Length 1444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTATGATCTCAATATACATGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG 60
DB 1 ATTATGATCTCAATATACATGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG 60
QY 61 AAAGTTTTCACGAATATGNACGACAAAATGGCTACCTCGATGTAATTTGGTATCTCAAC 120
DB 61 AAAGTTTTCACGAATATGNACGACAAAATGGCTACCTCGATGTAATTTGGTATCTCAAC 120
QY 121 TCAACATTATCTATPACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTATGTAAT 180
DB 121 TCAACATTATCTATPACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTATGTAAT 180
QY 181 GCAAGATCAGCATATGTAATTAATGTAATTCAGAAATCGTTTGGACGATTCGGATGTAAG 240
DB 181 GCAAGATCAGCATATGTAATTAATGTAATTCAGAAATCGTTTGGACGATTCGGATGTAAG 240
QY 241 TAGCCATTATTTAAATGTAATCTAATCTGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 300
DB 241 TAGCCATTATTTAAATGTAATCTAATCTGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 300
QY 301 ATTGTATAAATATCCATAAACACATCATGAAAGACATTTTCTTCAGGCTCTGAATTAAT 360
DB 301 ATTGTATAAATATCCATAAACACATCATGAAAGACATTTTCTTCAGGCTCTGAATTAAT 360
QY 361 TATGATACAAATTTCTAATAGAAAACGAATTAATTTAGTTGTAATGTAATGTAATGTAATGTAAT 420
DB 361 TATGATACAAATTTCTAATAGAAAACGAATTAATTTAGTTGTAATGTAATGTAATGTAATGTAAT 420
QY 421 AACAGCCCAACACGACGAGGACTAAGTTGCTGATGTAATGTAATGTAATGTAATGTAATGTAAT 480
DB 421 AACAGCCCAACACGACGAGGACTAAGTTGCTGATGTAATGTAATGTAATGTAATGTAATGTAAT 480
QY 481 TAAAAAACCGAGCTGTCTATGTAACACGCGGATCGAGCAGGTACACAGTCATGAAGCCATC 540
DB 481 TAAAAAACCGAGCTGTCTATGTAACACGCGGATCGAGCAGGTACACAGTCATGAAGCCATC 540
QY 541 AAAGCAAAAGAACTAATCCAAAGGGTGAGATGTAATTAATTAATTAATTAATTAATTAATTAAT 600
DB 541 AAAGCAAAAGAACTAATCCAAAGGGTGAGATGTAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 601 GAGGGAAGAGCTGTCTGACAGCAGGTACGTTATCTTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 GAGGGAAGAGCTGTCTGACAGCAGGTACGTTATCTTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 TGTCTGCT 720
DB 661 TGTCTGCT 720
QY 721 CGCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 CGCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 TTTTCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAGAAACGAAGAACTAA 840
DB 781 TTTTCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAGAAACGAAGAACTAA 840
QY 841 GAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTC 900
DB 841 GAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTC 900

DB 841 GAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTC 900
QY 901 ATCTTCTTCTCCGCTCTTTTCTTCCAAAGGTAATAGGAACCTTCTGGAATCTACTTTATTT 960
DB 901 ATCTTCTTCTCCGCTCTTTTCTTCCAAAGGTAATAGGAACCTTCTGGAATCTACTTTATTT 960
QY 961 GCTGATCTCGATCTGTTGTTTCTCAATTTCTTCTGAGATCTGGAATTCGTTAAATTTGGAT 1020
DB 961 GCTGATCTCGATCTGTTGTTTCTCAATTTCTTCTGAGATCTGGAATTCGTTAAATTTGGAT 1020
QY 1021 CTGGAACCTCACATAAATCTTTTGGTTTCTAGATCTGATCTGATCTGATCTGATCTGATCTGAT 1080
DB 1021 CTGGAACCTCACATAAATCTTTTGGTTTCTAGATCTGATCTGATCTGATCTGATCTGATCTGAT 1080
QY 1081 TAGCTCGAATATAGCTACCAAGATTTGGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1140
DB 1081 TAGCTCGAATATAGCTACCAAGATTTGGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1140
QY 1141 TACCTGGGAATGATTTGTATATGTAATTTGAAATCTGAACTCTGAACTCTGAACTCTGAACTCTGAA 1200
DB 1141 TACCTGGGAATGATTTGTATATGTAATTTGAAATCTGAACTCTGAACTCTGAACTCTGAACTCTGAA 1200
QY 1201 CTGAACATCTCAATGTTAGATTTGAATCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCTGAA 1260
DB 1201 CTGAACATCTCAATGTTAGATTTGAATCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCTGAA 1260
QY 1261 ATAGATCTTCTGAAACCTTAGATTTGTAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1320
DB 1261 ATAGATCTTCTGAAACCTTAGATTTGTAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1320
QY 1321 ATTCATCAGGCTTTATTTGATCTGATTTGAACTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1380
DB 1321 ATTCATCAGGCTTTATTTGATCTGATTTGAACTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1380
QY 1381 AATGCTCAGGCTGACGATATTAACCAATCGTGTGTCACAAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
DB 1381 AATGCTCAGGCTGACGATATTAACCAATCGTGTGTCACAAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
QY 1441 ATCC 1444
DB 1441 ATCC 1444
RESULT 2
ABK89341
ID ABK89341 standard; DNA; 6385 BP.
XX
AC ABK89341;
XX
DT 21-OCT-2002 (first entry)
XX
DE Plasmid pACT2B-GUS DNA.
XX
KW Arsenate reductase; ArsC; cyclic; circular; ds; antimionate; zinc;
KW cadmium; phytochelatase; biosynthetic enzyme; arsenite; cobalt;
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;
KW industrial waste; phytoremediation.
XX
OS Synthetic.
XX
PN WO200248335-A2.
XX
PD 20-JUN-2002.
XX
PF 13-DEC-2001; 2001WO-US048105.
XX
PR 13-DEC-2000; 2000US-0255001P.
XX
PR 22-JUN-2001; 2001US-0300525P.
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
PI Meagher RB, Li Y;
XX

DR WPI; 2002-593507/62.

XX Recombinant DNA molecules for producing transgenic plants tolerant to

PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of

PT contaminated soil or water, encodes arsenate reductase coding sequence.

XX Disclosure; Fig 13B; 131pp; English.

CC The invention relates to a nucleic acid molecule comprising a portion

CC encoding an arsenate reductase coding sequence and a plant-expressible

CC transcription regulatory sequence, the coding sequence being operably

CC linked to the transcription regulatory sequence. The nucleic acid is

CC useful for producing a plant which is resistant to at least one metal ion

CC such as a divalent cadmium ion or antimonate, by introducing the nucleic

CC acid molecule into a plant cell or into plant tissue, selecting for the

CC presence of the nucleic acid molecule to produce a transgenic plant cell

CC or plant tissue and regenerating a plant from the transgenic plant cell

CC or plant tissue. The method further comprises introducing at least one

CC plant-expressible phytochelatin biosynthetic enzyme coding sequence into

CC the cell to produce an arsenate metal ion resistant plant. The arsenate

CC reductase coding sequence is expressed under the control of a plant

CC promoter which directs expression in the above ground plant part. The

CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,

CC antimony, arsenate and arsenite ions. The transgenic plant is useful for

CC bioremediation of arsenate and/or cadmium-contaminated environments,

CC including soil, sediments, mine tailings, water, industrial waste,

CC groundwater and air. The transgenics are also useful for revegetation of

CC soils contaminated with metal ions and for removing and sequestering

CC these ions from water, wastewater and aqueous environments. The plants

CC are also useful for phytoremediation of contaminated soil, sediment,

CC water and mine tailings. This sequence represents plasmid pACT2B-GUS DNA,

CC used in the scope of the invention

XX

SQ Sequence 6385 BP; 1684 A; 1447 C; 1518 G; 1733 T; 0 U; 3 Other;

Query Match 92.9%; Score 1341.8; DB 6; Length 6385;

Best Local Similarity 98.8%; Pred. No. 5.8e-314; Indels 4; Gaps 2;

Matches 1375; Conservative 0; Mismatches 12;

QY 1 ATTATGATCTCAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTATGTAAG 60

DB 257 ATTATGATCTCAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAG 316

QY 61 AAAGTTTGAAGCAATATGNNACGACAAAATGCTACACTCGATGTAATGGTATCTCAAC 120

DB 317 AAAGTTTGAAGCAATATGNNACGACAAAATGCTAGACTCGATGTAATGGTATCTCAAC 376

QY 121 TCAACATATATCTTATACCAACATAGTTAGCAAAATTTAAACAACTATTTTATGTAT 180

DB 377 TCAACATATATCTTATACCAACATAGTTAGCAAAATTTAAACAACTATTTTATGTAT 436

QY 181 GCAAGAGTCAGCATATGTATATATTCAGATTCAGATCGTTTTCAGAGTTCGGATGTAGTAG 240

DB 437 GCAAGAGTCAGCATATGTATATATTCAGATTCAGATCGTTTTCAGAGTTCGGATGTAGTAG 496

QY 241 TAGCCATTATTTAATGTACATACTAATCGTGAATAGTATGATGAAACATTTGATCTTT 300

DB 497 TAGCCATTATTTAATGTACATACTAATCGTGAATAGTATGATGAAACATTTGATCTTT 556

QY 301 ATTGTATATAATTCATTAACACATCATGAAAGACACTTTCTTTCAGGGTCTGAATTAAT 360

DB 557 ATTGTATATAATTCATTAACACATCATGAAAGACACTTTCTTTCAGGGTCTGAATTAAT 616

QY 361 TATGATACAAATTTCTAATAGAAAACGAATTAATATTCGTTTGAATTCGTATGAATCTAAATTG 420

DB 617 TATGATACAAATTTCTAATAGAAAACGAATTAATATTCGTTTGAATTCGTATGAATCTAAATTG 676

QY 421 AACAGCCCAACCAACGACGAGGACTAACCGTTGCTCGATTTGACTCGGTTTAAAGTTAACCAAC 480

DB 677 AACAGCCCAACCAACGACGAGGACTAACCGTTGCTCGATTTGACTCGGTTTAAAGTTAACCAAC 736

QY 481 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATTCAGACAGGTTCACAGTCATGAAGCCATC 540

DB 737 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATTCAGACAGGTTCACAGTCATCAAGCCATC 796

QY 541 AAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTTTAAAAATTAGTTAAACAC 600

DB 797 AAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTTTAAAAATTAGTTAAACAC 856

QY 601 AAGGGAAAAA-GCTGTCTGACAGCCAGGTCACTTATCTTTTACCTGTGTGGTGGAAATGATTC 659

DB 857 AAGGGAAAAAAGGTGTCTGACAGCCAGGTCACTTATCTTTTACCTGTGTGGTGGAAATGATTC 916

QY 660 GTGTCTGTCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAAAGTTGTAAGAGATAAAC 719

DB 917 GTGTCTGTCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAAAGTTGTAAGAGATAAAC 976

QY 720 CGCCTATATATAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGCTCCTCTCA 779

DB 977 CGCCTATATATAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGCTCCTCTCA 1036

QY 780 CTTTCATCAGCGGTTTGAATTTCTCCGCGACTTTGACAGAGAAGAACAGAGAAAGACTA 839

DB 1037 CTTTCATCAGCGGTTTGAATTTCTCCGCGACTTTGACAGAGAAGAACAGAGAAAGACTA 1096

QY 840 AGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATTTCTCCTCAATCT 899

DB 1097 AGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATTTCTCCTCAATCT 1156

QY 900 CATCTTCTTTCTCCGCTCTTTCTTTTCAAGGTAATAGGAACCTTTCTGGATCTACTTTATT 959

DB 1157 CA---TCCTTCTCCGCTCTTTCTTTTCAAGGTAATAGGAACCTTTCTGGATCTACTTTATT 1213

QY 960 TGTGGAATCTGATCTTGTCTTCAATTTCTTCTGAGATCTGGAATTCGTTTAAATTTGA 1019

DB 1214 TGTGGAATCTGATCTTGTCTTCAATTTCTTCTGAGATCTGGAATTCGTTTAAATTTGA 1273

QY 1020 TCTGTGAACCTCCACTAATCTTTTGTGTTTACTAGATCGATCTAAGTTGACCGATCAG 1079

DB 1274 TCTGTGAACCTCCACTAATCTTTTGTGTTTACTAGATCGATCTAAGTTGACCGATCAG 1333

QY 1080 TTAGCTCGATTTATAGCTACAGAAATTTGGCTTTGACCTTGAGGAGAGATCCATGTTTCATG 1139

DB 1334 TTAGCTCGATTTATAGCTACAGAAATTTGGCTTTGACCTTGAGGAGAGATCCATGTTTCATG 1393

QY 1140 TTACTCGGAATGATTTGTATATGTGAATTTGAATCTGAATCTGTAAGTTGAGTTAGAA 1199

DB 1394 TTACTCGGAATGATTTGTATATGTGAATTTGAATCTGAATCTGTAAGTTGAGTTAGAA 1453

QY 1200 TCTGAACTCTCAATGTAGATTTGAATCTGAACTGTAAGTTAGTTAGTTAGTTAGTTAG 1259

DB 1454 TCTGAACTCTCAATGTAGATTTGAATCTGAACTGTAAGTTAGTTAGTTAGTTAGTTAG 1513

QY 1260 TATAGATTTCTTCGAAACCTTAGGATTTTGTAGTGTCTGATGTTGAACAGAAAGCTATTTCT 1319

DB 1514 TATAGATTTCTTCGAAACCTTAGGATTTTGTAGTGTCTGATGTTGAACAGAAAGCTATTTCT 1573

QY 1320 GATTCAATCAGGGTTTATTTGACTGTATTTGAATCTTTTGTGTTGTTGAGTCTCATAAA 1379

DB 1574 GATTCAATCAGGGTTTATTTGACTGTATTTGAATCTTTTGTGTTGTTGAGTCTCATAAA 1633

QY 1380 AAATGGCTGAG 1390

DB 1634 GGATCCCGGG 1644

RESULT 3

ID ABR89340 standard; DNA; 4526 BP.

XX

AC ABR89340;

XX

DT 21-OCT-2002 (first entry)

XX

DE Plasmid pACT2B DNA.

XX

KW Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;
KW cadmium; phytochelatin biosynthetic enzyme; arsenite; arsenate; cobalt;
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;
KW industrial waste; phytoremediation.
XX
OS Synthetic.
XX
XX WO200248335-A2.
XX
XX 20-JUN-2002.
XX
XX 13-DEC-2001; 2001WO-US048105.
XX
XX 13-DEC-2000; 2000US-0255001P.
PR 22-JUN-2001; 2001US-0300525P.
XX
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
PA
XX Meagher RB, Li Y;
PI
XX WPI; 2002-583507/62.
DR
XX Recombinant DNA molecules for producing transgenic plants tolerant to
PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of
PT contaminated soil or water, encodes arsenate reductase coding sequence.
XX
XX Disclosure; Fig 12B; 131pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a portion
CC encoding an arsenate reductase coding sequence and a plant-expressible
CC transcription regulatory sequence, the coding sequence being operably
CC linked to the transcription regulatory sequence. The nucleic acid is
CC useful for producing a plant which is resistant to at least one metal ion
CC such as a divalent cadmium ion or antimonate, by introducing the nucleic
CC acid molecule into a plant cell or into plant tissue, selecting for the
CC presence of the nucleic acid molecule to produce a transgenic plant cell
CC or plant tissue and regenerating a plant from the transgenic plant cell
CC or plant tissue. The method further comprises introducing at least one
CC plant expressible phytochelatin biosynthetic enzyme coding sequence into
CC the cell to produce an arsenate metal ion resistant plant. The arsenate
CC reductase coding sequence is expressed under the control of a plant
CC promoter which directs expression in the above ground plant part. The
CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,
CC antimony, arsenate and arsenite ions. The transgenic plant is useful for
CC bioremediation of arsenate and/or cadmium-contaminated environments,
CC including soil, sediments, mine tailings, water, industrial waste,
CC groundwater and air. The transgenics are also useful for revegetation of
CC soils contaminated with metal ions and for removing and sequestering
CC these ions from water, wastewater and aqueous environments. The plants
CC are also useful for phytoremediation of contaminated soil, sediment,
CC water and mine tailings. This sequence represents plasmid pACT2B DNA,
CC used in the scope of the invention
XX
SQ Sequence 4526 BP; 1203 A; 997 C; 1002 G; 1321 T; 0 U; 3 Other;

Query Match 92.9%; Score 1341.4; DB 6; Length 4526;
Best Local Similarity 98.9%; Pred. No. 5.8e-314;
Matches 13/4; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

QY 1 ATTATGATCTCAATAATGATGATACATATCTCATCTAGATCTAGGTTATCATTTATGTTAAG 60
DB |||||
DB 257 ATTATGATCTCAATAATGATGATACATATCTCATCTAGATCTAGGTTATCATTTATGTTAAG 316
QY 61 AAAGTTTTCAGCAATATGNNACGACAAATGGCTACACTCGATGTAAATGGTATCTCAAC 120
DB |||||
DB 317 AAAGTTTTCAGCAATATGNNACGACAAATGGCTACACTCGATGTAAATGGTATCTCAAC 376
QY 121 TCACATTTATCTTATACCAACATTTAGTACCAAAATTTAAACACTATTTTATGTTAT 180
DB |||||
DB 377 TCACATTTATCTTATACCAACATTTAGTACCAAAATTTAAACACTATTTTATGTTAT 436
QY 181 GCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTGACGAGTTCGAGTGTAGTAG 240
DB |||||

DB 437 GCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTGACGAGTTCGAGTGTAGTAG 496
QY 241 TAGCCATTATTTAAATGATACATACATAATCGTGAATAGTGTATGATGAACAACATTGTATCTT 300
DB |||||
DB 497 TAGCCATTATTTAAATGATACATACATAATCGTGAATAGTGTATGATGAACAACATTGTATCTT 556
QY 301 ATTGTATAAATATCCATAAAACACATCATGAAAGACACATTTCTTTCCAGGTCCTGAATTAAT 360
DB |||||
DB 557 ATTGTATAAATATCCATAAAACACATCATGAAAGACACATTTCTTTCCAGGTCCTGAATTAAT 616
QY 361 TATGATACAAATTTAAATGAGAAAAACGAATTAATTTAGTTGATGATGATGAATCTTAATTTG 420
DB |||||
DB 617 TATGATACAAATTTAAATGAGAAAAACGAATTAATTTAGTTGATGATGATGAATCTTAATTTG 676
QY 421 AACAGCCCAACACGACGAGGACTAAACGTTGCTCGATTTGACTCGGTTTAAAGTTAAACCAAC 480
DB |||||
DB 677 AACAGCCCAACACGACGAGGACTAAACGTTGCTCGATTTGACTCGGTTTAAAGTTAAACCAAC 736
QY 481 TAAAAAAGCGAGCTGTCTCATGTATAACACGCGGATCGAGCAGGTCACAGTCATGAAGCCATC 540
DB |||||
DB 737 TAAAAAAGCGAGCTGTCTCATGTATAACACGCGGATCGAGCAGGTCACAGTCATGAAGCCATC 796
QY 541 AAAGCAAAAGAACTAATCCAAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAAACAC 600
DB |||||
DB 797 AAAGCAAAAGAACTAATCCAAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAAACAC 856
QY 601 GAGGGAAA-GCTGTCCTGACAGCCAGGTCAGTTATCTTTACCTCTGTGTGCGAAATGATTC 659
DB |||||
DB 857 GAGGGAAAAGGCTGTCTGACAGCCAGGTCAGTTATCTTTACCTGTGTGCGAAATGATTC 916
QY 660 GTGCTGTGCGATTTTAAATTAATTTTAAAAAGCGCGGAAAAATAAAGTTGTGAAGAGATAAAC 719
DB |||||
DB 917 GTGCTGTGCGATTTTAAATTAATTTTAAAAAGCGCGGAAAAATAAAGTTGTGAAGAGATAAAC 976
QY 720 CGCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGCTCGTTGTCCTCTCA 779
DB |||||
DB 977 CGCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGCTCGTTGTCCTCTCA 1036
QY 780 CTTTCATCAGCGTTTTCGAAATCTCCGGGACTTGACAGAGAAACAAAGGAAGAAGACTA 839
DB |||||
DB 1037 CTTTCATCAGCGTTTTCGAAATCTCCGGGACTTGACAGAGAAACAAAGGAAGAAGACTA 1096
QY 840 AGAGAAAAGTAAGAGATAAATCCAGGAGATTCATTTCCGTTTGAATCTTCTCCTCAATCT 899
DB |||||
DB 1097 AGAGAAAAGTAAGAGATAAATCCAGGAGATTCATTTCCGTTTGAATCTTCTCCTCAATCT 1156
QY 900 CATCTCTTCTTCCGCTCTTCTTTTCCAAAGTAATAGAACTTTCTCGATCTACTTTATT 959
DB |||||
DB 1157 CA---TCTTCTTCCGCTCTTCTTTTCCAAAGGTAATAGGAACCTTCTGGATCTACTTTATT 1213
QY 960 TGCTCGATCTCGATCTTGTGTTTCTCAATTTCTCAGATCTGGAATTCGTTTAAATTTGGA 1019
DB |||||
DB 1214 TGCTCGATCTCGATCTTGTGTTTCTCAATTTCTTGAGATCTGGAATTCGTTTAAATTTGGA 1273
QY 1020 TCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAATTCGATCTAAGTTGACCGATCAG 1079
DB |||||
DB 1274 TCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAATTCGATCTAAGTTGACCGATCAG 1333
QY 1080 TTAGTCGATTTATAGTACACAGAAATTTGGCTTTGACCTTGATGGAGAGATCCATGTTTCATG 1139
DB |||||
DB 1334 TTAGTCGATTTATAGTACACAGAAATTTGGCTTTGACCTTTGATGGAGAGATCCATGTTTCATG 1393
QY 1140 TTACCTCGGAATGATTTGTATGTAATCTGAATTCGAAATCTGAATTTGAAGTTAGATTGAA 1199
DB |||||
DB 1394 TTACCTCGGAATGATTTGTATGTAATTCGAAATCTGAATTTGAAGTTAGATTGAA 1453
QY 1200 TCTGAACACTGTCTCAATTTAGATTGAATCTGAACACTCTGTTTAAAGTTAGATGAAGTTTGTG 1259
DB |||||
DB 1454 TCTGAACACTGTCTCAATTTAGATTGAATCTGAACACTCTGTTTAAAGTTAGATGAAGTTTGTG 1513
QY 1260 TATAGATTTCTCGAAACCTTAGGATTTGATGTCGTAGCTTGTGAACAGAAAGCTATTTCCT 1319
DB |||||
DB 1514 TATAGATTTCTCGAAACCTTTAGGATTTGATGTCGTAGCTTGTGAACAGAAAGCTATTTCCT 1573

Db 10700 AACCCGCTATATAAATTCATATATTTCTCTCCGCTTTGAAATGTCTCGTGTCTCTCC 10759
Qy 777 TCACCTTTTCATAGCCGCTTTTGAATCTCCGCGCATTTGACAGAGAAAGAAAGAAAGA 836
Db 10760 TCACCTTTTCATAGCCGCTTTTGAATCTCCGCGCATTTGACAGAGAAAGAAAGAAAGA 10819
Qy 837 CTAAGAGAGAAAGTAAGAGATATCCAGAGATTCATCTCCGCTTTGAAATCTCTCTCAA 896
Db 10820 CTAAGAGAGAAAGTAAGAGATATCCAGAGATTCATCTCCGCTTTGAAATCTCTCTCAA 10879
Qy 897 TCTCATCT 956
Db 10880 TCTCA---TCT 10936
Qy 957 ATTGCTGAGATCGATCT 1016
Db 10937 ATTGCTGAGATCGATCT 10996
Qy 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGAT 1076
Db 10997 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGAT 11056
Qy 1077 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTTGACCTTCATCGAGAGATCCCATGTTTC 1136
Db 11057 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTTGACCTTCATCGAGAGATCCCATGTTTC 11116
Qy 1137 ATGTTACCTGGGAAATGATTTTGTATATGTGAATGAAATCTGAACTGTTGAAATTTAGATT 1196
Db 11117 ATGTTACCTGGGAAATGATTTTGTATATGTGAATGAAATCTGAACTGTTGAAATTTAGATT 11176
Qy 1197 GAATCTGAACACTGTCATGTTAGATTTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255
Db 11177 GAATCTGAACACTGTCATGTTAGATTTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT 11236
Qy 1256 TGTGTATAGATTTCTTGGAAACCTTAGGATTTCTGTAGTGTGCTAGCTTGAACAGAAAGCTAT 1315
Db 11237 TGTGTATAGATTTCTTGGAAACCTTAGGATTTCTGTAGTGTGCTAGCTTGAACAGAAAGCTAT 11296
Qy 1316 TTCTGATTCATTCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTGTGTTTGCAGCTCA 1375
Db 11297 TTCTGATTCATTCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTGTGTTTGCAGCTCA 11356
Qy 1376 TAAAAAATGGCTGAGGCTGACGATAT 1401
Db 11357 TAAAAAATGGCTGAGGCTGACGATAT 11382

RESULT 6

ABK89343
ID ABK89343 standard; DNA; 3450 BP.
XX
AC ABK89343;
XX
XX
DT 21-OCT-2002 (first entry)
XX
DE Plasmid pBCSACT2 DNA.
XX
XX Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;
KW cadmium; phytochelatin biosynthetic enzyme; arsenate; arsenite; cobalt;
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;
KW industrial waste; phytoremediation.
XX
OS Synthetic.
XX
XX WO200248335-A2.
XX
PD 20-JUN-2002.
XX
XX 13-DEC-2001; 2001WO-US048105.
PF
XX 13-DEC-2000; 2000US-0255001P.
PR
XX 22-JUN-2001; 2001US-0300525P.
XX

(UYGE-) UNIV GEORGIA RES FOUND INC.
Meagher RB, Li Y;
WPI; 2002-583507/62.
Recombinant DNA molecules for producing transgenic plants tolerant to heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of contaminated soil or water, encodes arsenate reductase coding sequence.
Disclosure; Page 75-77; 131pp; English.
The invention relates to a nucleic acid molecule comprising a portion encoding an arsenate reductase coding sequence and a plant-expressible transcription regulatory sequence, the coding sequence being operably linked to the transcription regulatory sequence. The nucleic acid is useful for producing a plant which is resistant to at least one metal ion such as a divalent cadmium ion or antimonate, by introducing the nucleic acid molecule into a plant cell or into plant tissue, selecting for the presence of the nucleic acid molecule to produce a transgenic plant cell or plant tissue and regenerating a plant from the transgenic plant cell or plant tissue. The method further comprises introducing at least one plant-expressible phytochelatin biosynthetic enzyme coding sequence into the cell to produce an arsenate metal ion resistant plant. The arsenate reductase coding sequence is expressed under the control of a plant promoter which directs expression in the above ground plant part. The plant is also resistant to cadmium, cobalt, copper, mercury, zinc, antimony, arsenate and arsenite ions. The transgenic plant is useful for bioremediation of arsenate and/or cadmium-contaminated environments, including soil, sediments, mine tailings, water, industrial waste, groundwater and air. The transgenics are also useful for revegetation of soils contaminated with metal ions and for removing and sequestering these ions from water, wastewater and aqueous environments. The plants are also useful for phytoremediation of contaminated soil, sediment, water and mine tailings. This sequence represents plasmid pBCSACT2 DNA, used in the scope of the invention
Sequence 3450 BP; 946 A; 694 C; 764 G; 1046 T; 0 U; 0 Other;
Query Match 89.9%; Score 1298.4; DB 6; Length 3450;
Best Local Similarity 98.1%; Pred. No. 1.4e-303;
Matches 1376; Conservative 0; Mismatches 18; Indels 8; Gaps 6;
Qy 1 ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG 60
Db 9 ATGCTGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG 68
Qy 61 AAAGTTTGGCAATATGNNACGACAAATGGCTACACTCGATGTAATGGTATCTCAAC 120
Db 69 AAAGTTTGGCAATATGGCAGCAGCAAAATGGCTAGACTCGATGTAATGGTATCTCAAC 128
Qy 121 TCAACATTATATCTTATCCAAACATTAGTTAG-CAAAATTTAAACAACA-TTTTTATGT 178
Db 129 TCAACATTATATCTTATCCAAACATTAGTTAGACAAAATTTAAACAACA-TTTTTATGT 188
Qy 179 ATGCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTGGACGAGTTCGGATGTAGT 238
Db 189 ATGCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTGGACGAGTTCGGATGTAGT 248
Qy 239 AGTAGCCATTTATTTAATGTACATCTAATCTGTGAATAGTG-ATATGATGAACAATTTGAT 297
Db 249 AGTAGCCATTTATTTAATGTACATCTAATCTGTGAATAGTGAAATATGATGAACAATTTGAT 308
Qy 298 CTTATTTGATAAATATCCATAAACACATCATGAAGACACCTTTCTTTCAGGCTCTGAATT 357
Db 309 CTTATTTGATAAATATCCATAAACACATCATGAAGACACCTTTCTTTCAGGCTCTGAATT 368
Qy 358 AATTATGATACAAATTTCTAATAGAAAAAGAAATTAATTAATACGTTGAATTTGATGAATCTAA 417
Db 369 AATTATGATACAAATTTCTAATAGAAAAAGAAATTAATTAATACGTTGAATTTGATGAATCTAA 428
Qy 418 TTGAACAGCCAAACCAAGAGAGACTAAACGTTCCCTGGATTTGACTCGGTTTAAGTTAAC 477

Db 429 TTGAACAAGCCAAACACGACGACGACTAACGTTGCTGGAATGACTCGGTTAAGTTAAC 488
Qy 478 CACTAAAAACGGAGCTGCTATGTAACACGCGGATCGAGGAGGTCACAGTCATGAAGCC 537
Db 489 CACTAAAAACGGAGCTGCTATGTAACACGCGGATCGAGGAGGTCACAGTCATGAAGCC 548
Qy 538 ATCAAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAAGTTAAAAATTAAGTTAA 597
Db 549 ATCAAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAAGTTAAAAATTAAGTTAA 608
Qy 598 CACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACTGTTATCTTTACTGTTGTCGAAATGA 656
Db 609 CACGAGGAAAAAGGCTGCTGACAGCCAGGTCACTGTTATCTTTACTGTTGTCGAAATGA 668
Qy 657 TTTCTGTCTGCTCGATTTTAAATTAATTTTAAAGGCGGAAATAAAGTTGTAAGAGATA 716
Db 669 TTTCTGTCTGCTCGATTTTAAATTAATTTTAAAGGCGGAAATAAAGTTGTAAGAGATA 728
Qy 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGCTCTGTTGCTCTCC 776
Db 729 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGCTCTGTTGCTCTCC 788
Qy 777 TCACCTTTCATAGCCGCTTTTGAATCTCCGCGACCTTGACAGAGAAACAAGAAAGAGA 836
Db 789 TCACCTTTCATAGCCGCTTTTGAATCTCCGCGACCTTGACAGAGAAACAAGAAAGAGA 848
Qy 837 CTAAGAGAGAAAGTAAGAGATTAATCCAGGAGATTCATTTCTCGTTTGAATCTTCTCAA 896
Db 849 CTAAGAGAGAAAGTAAGAGATTAATCCAGGAGATTCATTTCTCGTTTGAATCTTCTCAA 908
Qy 897 TCTCATCTTCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTTCGGAATCTACTTT 956
Db 909 TCTCA---TCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTTCGGAATCTACTTT 965
Qy 957 ATTGCTGAGCTCGATCTGTTTCTCAATTTCTTGAGATCTGGAATCTGTTAAATTT 1016
Db 966 ATTGCTGAGCTCGATCTGTTTCTCAATTTCTTGAGATCTGGAATCTGTTAAATTT 1025
Qy 1017 GATCTGTGAACCTCCACTAAATCTTTGTTTACTAGAACTGAATCGATCTAAGTTGACCGAT 1076
Db 1026 GATCTGTGAACCTCCACTAAATCTTTGTTTACTAGAACTGAATCGATCTAAGTTGACCGAT 1085
Qy 1077 CAGTTAGCTCGATTTAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCCATGTTTC 1136
Db 1086 CAGTTAGCTCGATTTAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCCATGTTTC 1145
Qy 1137 ATGTTACCTGGGAAATGATTTGATATGTAATGAAATCTGAAATCTGTTGAAAGTTAGATT 1196
Db 1146 ATGTTACCTGGGAAATGATTTGATATGTAATGAAATCTGAAATCTGTTGAAAGTTAGATT 1205
Qy 1197 GAATCTGAACCTGTCATGTTAGATTGAATCTGACACCTGTTTAA-GTTAGATGAAGTT 1255
Db 1206 GAATCTGAACCTGTCATGTTAGATTGAATCTGACACCTGTTTAAAGTTAGATGAAGTT 1265
Qy 1256 TGTGTATAGATTCTTCCGAAACCTTAGGATTTGTAGTGTCTAGCTTGAACAGAAAGCTAT 1315
Db 1266 TGTGTATAGATTCTTCCGAAACCTTAGGATTTGTAGTGTCTAGCTTGAACAGAAAGCTAT 1325
Qy 1316 TTTCTGATCAATCAGGGTTTAAATTTGATCTGTAATGAACCTCTTTTGTGTTTGCAGCTCA 1375
Db 1326 TTTCTGATCAATCAGGGTTTAAATTTGATCTGTAATGAACCTCTTTTGTGTTTGCAGCTCA 1385
Qy 1376 TAAAAAATGGCTGAGGCTGACG 1397
Db 1386 TAAACCATGGCAATCCCGGACG 1407

RESULT 7
ADN61599
ID ADN61599 standard; DNA; 7474 BP.
XX
AC
ADN61599;
XX

DT 29-JUL-2004 (first entry)
XX COT102 event nucleotide sequence SEQ ID NO:7.
XX COT102 motif event; insect resistant plant; VIP3A; detection; plant;
KW COT102 event; insecticidal; Heliothis; Helicoverpa; Spodoptera;
KW cotton boll worm; gene; ds; insect resistant transgenic cotton event.
OS Synthetic.
XX WO2004039986-A1.
XX 13-MAY-2004.
XX 23-OCT-2003; 2003WO-EP011725.
XX 29-OCT-2002; 2002GB-00025129.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA Ellis DM, Negrotto DV, Shi L, Shotkoski FA, Thomas CR;
PI WPI; 2004-390327/36.
DR Novel polynucleotide comprising contiguous nucleotides of motif
PT designated COT102 event, useful in detecting insect resistant plant
PT material derived from the event.
XX Claim 16; SEQ ID NO 7; 64pp; English.
XX The present invention describes a polynucleotide (I) comprising at least
CC 17 contiguous nucleotide of the insect resistant transgenic cotton
CC (COT102) motif event having the 26 nucleotide sequence of SEQ ID NO:1 or
CC SEQ ID NO:2. Also described: (1) an insect resistant plant (II)
CC comprising a VIP3A protein and (I); (2) detecting (M1) plant material
CC derived from the COT102 event; and (3) a kit of parts comprising a unit
CC for (M1). (I) and (M1) are useful for detecting a plant material derived
CC from the COT102 event. (I) enables efficient detection of a plant
CC material derived from the COT102 event. (II) has an insecticidal effect
CC on insects from one or more species chosen from Heliothis sp.,
CC Helicoverpa sp. and Spodoptera sp.. (II) has enhanced self-defence
CC mechanism against infestation by pest insects such as Helicoverpa zea
CC (cotton boll worm), and so reduces the number of insecticide sprays
CC during the cultivation of (II) compared to non-transgenic cotton plant of
CC the same variety and yield loss through insect pests in kept at a minimal
CC level. The present sequence represents a COT102 event nucleotide
CC sequence, which is used in the exemplification of the present invention.
XX
SQ Sequence 7474 BP; 2223 A; 1777 C; 1647 G; 1827 T; 0 U; 0 Other;
Query Match 89.9%; Score 1298.4; DB 12; Length 7474;
Best Local Similarity 98.8%; Pred. No. 1.7e-303;
Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;
Qy 1 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGTTATCATTTATGTAAG 60
Db 3228 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGTTATCATTTATGTAAG 3287
Qy 61 AAAGTTTTGACGAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
Db 3288 AAAGTTTTGACGAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 3347
Qy 121 TCAACATTATCTTATATACCAACATTAGTAG-CAAAAATTTAAACAACATA-TTTTATATGT 178
Db 3348 TCAACATTATCTTATATACCAACATTAGTAGAGAAAATTTAAACAACATA-TTTTATATGT 3407
Qy 179 ATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTCAGAGTTCGGATGTAGT 238
Db 3408 ATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTCAGAGTTCGGATGTAGT 3467
Qy 239 AGTAGCCATTATTTAATGTACATCTAATCTGTAATAGTG-ATATGATGAACATTTAT 297
Db 3468 AGTAGCCATTATTTAATGTACATCTAATCTGTAATAGTGATGTAATGATGAACATTTAT 3527

QY 298 CTTATTGTATTAATATCCATTAACACATCATGAAAGACACTTTCTTTTCAGGGTCTGAATT 357
DB |||||||
QY 3528 CTTATTGTATTAATATCCATTAACACATCATGAAAGACACTTTCTTTTCAGGGTCTGAATT 3587
DB |||||||
QY 358 AATTATTGATACAAATCTTAATAGAAAACGAATTAATTAAGTTGATGAAATCTTAA 417
DB |||||||
QY 3588 AATTATTGATACAAATCTTAATAGAAAACGAATTAATTAAGTTGATGAAATCTTAA 3647
DB |||||||
QY 418 TTGAACAAGCCCAACACGACGAGGACTAAGCTTGCCTGGATTGACTCGGTTTAAAGTTAA 477
DB |||||||
QY 3648 TTGAACAAGCCCAACACGACGAGGACTAAGCTTGCCTGGATTGACTCGGTTTAAAGTTAA 3707
DB |||||||
QY 478 CACTTAATAAAGCGAGCTGCTATGTAACACCGCGGATCGAGAGGTCAAGTCATGAAGCC 537
DB |||||||
QY 3708 CACTTAATAAAGCGAGCTGCTATGTAACACCGCGGATCGAGAGGTCAAGTCATGAAGCC 3767
DB |||||||
QY 538 ATCAAGCAAGCAAGCACTAATCAAGGGGTGAGATGATTAATTAAGTTTAAAGTTAA 597
DB |||||||
QY 3768 ATCAAGCAAGCAAGCACTAATCAAGGGGTGAGATGATTAATTAAGTTTAAAGTTAA 3827
DB |||||||
QY 598 CACGAGGGAAAA-GCTGTCTGACAGCCAGGTCAAGTTATCTTTACCTGTGGTCGAAATGA 656
DB |||||||
QY 3828 CACGAGGGAAAAAGGCTGTCTGACAGCCAGGTCAAGTTATCTTTACCTGTGGTCGAAATGA 3887
DB |||||||
QY 657 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGCCCGAAATAAAGTTGTAAGAGATA 716
DB |||||||
QY 3888 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGCCCGAAATAAAGTTGTAAGAGATA 3947
DB |||||||
QY 717 AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAAATTTGTCGTGTTGCTTCC 776
DB |||||||
QY 3948 AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAAATTTGTCGTGTTGCTTCC 4007
DB |||||||
QY 777 TCACCTTTTCATCAGCCGCTTTGAAATCTCCGCGACTTGACAGAGAAAGCAAGAAAGA 836
DB |||||||
QY 4008 TCACCTTTTCATCAGCCGCTTTGAAATCTCCGCGACTTGACAGAGAAAGCAAGAAAGA 4067
DB |||||||
QY 837 CTAAGAGAAAGTAAGAGATAATCAGGAGATCAATTCGCTTTGAAATCTTCTCTCAA 896
DB |||||||
QY 4068 CTAAGAGAAAGTAAGAGATAATCAGGAGATCAATTCGCTTTGAAATCTTCTCTCAA 4127
DB |||||||
QY 897 TCTCATCTCTCTCTCCGCTCTTTCTCTCCAGGTAATAGGAACCTTCTGGATCTACTTT 956
DB |||||||
QY 4128 TCTCA---TCTTCTCTCCGCTCTTTCTCTCCAGGTAATAGGAACCTTCTGGATCTACTTT 4184
DB |||||||
QY 957 ATTGCTGATCTGATCTGTTGTTTCTCAATTTCTCTGAGATCTGAAATTCGTTTAAATTT 1016
DB |||||||
QY 4185 ATTGCTGATCTGATCTGTTGTTTCTCAATTTCTCTGAGATCTGAAATTCGTTTAAATTT 4244
DB |||||||
QY 1017 GGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGAATCGATCTAAGTTGACCGAT 1076
DB |||||||
QY 4245 GGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGAATCGATCTAAGTTGACCGAT 4304
DB |||||||
QY 1077 CAGTTAGCTGATATAGTACAGAAATTTGCTTGAACCTTGAATCGAGAGATCCATGTT 1136
DB |||||||
QY 4305 CAGTTAGCTGATATAGTACAGAAATTTGCTTGAACCTTGAATCGAGAGATCCATGTT 4364
DB |||||||
QY 1137 ATGTTACCTCGGAATGATTTGATATGTAATCGAAATCTCAATCTGTTGAAATTTAGATT 1196
DB |||||||
QY 4365 ATGTTACCTCGGAATGATTTGATATGTAATCGAAATCTCAATCTGTTGAAATTTAGATT 4424
DB |||||||
QY 1197 GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255
DB |||||||
QY 4425 GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT 4484
DB |||||||
QY 1256 TGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGCTTGAACAGAAAGCTAT 1315
DB |||||||
QY 4485 TGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGCTTGAACAGAAAGCTAT 4544
DB |||||||
QY 1316 TTCGATTTCAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTCTGTGTTTTCAGCTCA 1375
DB |||||||
QY 4545 TTCGATTTCAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTCTGTGTTTTCAGCTCA 4604
DB |||||||

QY 1376 TAAAAA 1381
DB 4605 TAAAAA 4610

RESULT 8
ADN61613
ID ADN61613 standard; DNA; 9356 BP.

XX ADN61613;
XX 29-JUL-2004 (first entry)
XX

DE COT102 event nucleotide sequence SEQ ID NO:21.

XX COT102 motif event; insect resistant plant; VIP3A; detection; plant;
KW COT102 event; insecticidal; Heliothis; Helicoverpa; Spodoptera;
KW cotton boll worm; gene; ds; insect resistant transgenic cotton event.
XX Synthetic.

XX WO2004039986-A1.

PD 13-MAY-2004.

XX 23-OCT-2003; 2003WO-EP011725.

XX 29-OCT-2002; 2002GB-00025129.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Ellis DM, Negrotto DV, Shi L, Shotkoski FA, Thomas CR;

XX WPI; 2004-390327/36.

PT Novel polynucleotide comprising contiguous nucleotides of motif
PT designated COT102 event, useful in detecting insect resistant plant
PT material derived from the event.

PS Claim 9; SEQ ID NO 21; 64pp; English.

XX The present invention describes a polynucleotide (I) comprising at least
CC 17 contiguous nucleotide of the insect resistant transgenic cotton
CC (COT102) motif event having the 26 nucleotide sequence of SEQ ID NO:1 or
CC SEQ ID NO:2. Also described: (i) an insect resistant plant (II)
CC comprising a VIP3A protein and (I); (2) detecting (MI) plant material
CC derived from the COT102 event; and (3) a kit of parts comprising a unit
CC for (MI). (I) and (MI) are useful for detecting a plant material derived
CC from the COT102 event. (I) enables efficient detection of a plant
CC material derived from the COT102 event. (II) has an insecticidal effect
CC on insects from one or more species chosen from *Heliothis* sp.,
CC *Helicoverpa* sp. and *Spodoptera* sp.. (II) has enhanced self-defence
CC mechanism against infestation by pest insects such as *Helicoverpa* zea
CC (cotton boll worm), and so reduces the number of insecticide sprays
CC during the cultivation of (II) compared to non-transgenic cotton plant of
CC the same variety and yield loss through insect pests in kept at a minimal
CC level. The present sequence represents a COT102 event nucleotide
CC sequence, which is used in the exemplification of the present invention.

XX Sequence 9356 BP; 2869 A; 2138 C; 1953 G; 2396 T; 0 U; 0 Other;

Query Match 89.9%; Score 1298.4; DB 12; Length 9356;
Best Local Similarity 98.8%; Pred. NO. 1.9e-303;
Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAATATACATTTGATACATATCTCATCTAGATCTAGTTATCATTTATGTAAG 60
DB 4614 ATTATGATCTCAATATACATTTGATACATATCTCATCTAGATCTAGTTATCATTTATGTAAG 4673

QY 61 AAAGTTTTGACGAATATGNNACGAAATCGCTACACTCGATGTAATTTGGTATCTCAAC 120
DB 4674 AAAGTTTTGACGAATATGNNACGAAATCGCTAGACTCGATGTAATTTGGTATCTCAAC 4733

QY 121 TCAACATTATCTATTACCAAAACATTAGTAGT-CAAAAATTTAAACAACCTA-TTTTATGT 178
|||||
Db 4734 TCAACATTATCTATTACCAAAACATTAGTAGAGAAAATTTAAACAACCTATTTTATGT 4793
|||||
QY 179 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAAATCGTTTTCAGCAGATTCGGATGTAGT 238
|||||
Db 4794 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAAATCGTTTTCAGCAGATTCGGATGTAGT 4853
|||||
QY 239 AGTAGCCATTATTTAATGTACATCTAATTCGTTGAATAGTG-ATATGATGAACATTTGTAT 297
|||||
Db 4854 AGTAGCCATTATTTAATGTACATCTAATTCGTTGAATAGTGAAATATGATGAACATTTGTAT 4913
|||||
QY 298 CTTATTGTATAAATATCCATAAACAACATCATCAAAAGACACATTTCTTTTCAGGCTCTGAATT 357
|||||
Db 4914 CTTATTGTATAAATATCCATAAACAACATCAAAAGACACATTTCTTTTCAGGCTCTGAATT 4973
|||||
QY 358 AATTATGATCAATTTCTAATAGAAAACGAATTAATTTAGTTGAAATTTGATGAATCTAA 417
|||||
Db 4974 AATTATGATCAATTTCTAATAGAAAACGAATTTAATTTAGCTTTGAATTTGATGAATCTAA 5033
|||||
QY 418 TTGAACAGCCCAACACGAGGAGGACTAACGTTGCTCGATTGACTCGGTTTAACTTAAC 477
|||||
Db 5034 TTGAACAGCCCAACACGAGGAGGACTAACGTTGCTCGATTGACTCGGTTTAACTTAAC 5093
|||||
QY 478 CACTAAAAAAGCGAGCTCTCATGTAAACACGCGGATCGAGCAGGTCAACAGTCATCAAGCC 537
|||||
Db 5094 CACTAAAAAAGCGAGCTCTCATGTAAACACGCGGATCGAGCAGGTCAACAGTCATCAAGCC 5153
|||||
QY 538 ATCAAAGCAAAAGAACTAATCCAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAA 597
|||||
Db 5154 ATCAAAGCAAAAGAACTAATCCAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAA 5213
|||||
QY 598 CACGAGGGGAAA-GCTGCTGACAGCCAGGTCAAGTTATCTTTTACCTGCTGGTCGAAAATGA 656
|||||
Db 5214 CACGAGGGGAAAAGGGTGTCTGACAGCCAGGTCAAGTTATCTTTTACCTGCTGGTCGAAAATGA 5273
|||||
QY 657 TTCTGCTCTGCTCGATTTTAAATTAATTTTGTAAAGGCCGAAAATAAAGTTGTAAGAGATA 716
|||||
Db 5274 TTCTGCTCTGCTCGATTTTAAATTAATTTTGTAAAGGCCGAAAATAAAGTTGTAAGAGATA 5333
|||||
QY 717 AACCCGCCATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGTCGTTGCTCTCC 776
|||||
Db 5334 AACCCGCCATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGTCGTTGCTCTCC 5393
|||||
QY 777 TCACATTTCATCAGCCGTTTTGATCTCCGCGACCTTGACAGAGAAACAAGAAAGAGA 836
|||||
Db 5394 TCACATTTCATCAGCCGTTTTGATCTCCGCGACCTTGACAGAGAAACAAGAAAGAGA 5453
|||||
QY 837 CTAAGAGAGAAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTTTGAATCTTCTCTCAA 896
|||||
Db 5454 CTAAGAGAGAAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTTTGAATCTTCTCTCAA 5513
|||||
QY 897 TCTCATCTTTCTTCTCCGCTCTTTCTTTCCAAAGGTAATAGGAACCTTTCTGATCTACTTT 956
|||||
Db 5514 TCTCA--TCTTCTTCCGCTCTTTCTTTCCAAAGGTAATAGGAACCTTTCTGATCTACTTT 5570
|||||
QY 957 ATTTGCTGATCTCGATCTGTTTTCTCAATTTCTTGAGATCTGGAATTCGTTAAATTT 1016
|||||
Db 5571 ATTTGCTGATCTCGATCTGTTTTCTCAATTTCTTGAGATCTGGAATTCGTTAAATTT 5630
|||||
QY 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAATCGATCTAAGTTGACCGAT 1076
|||||
Db 5631 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAATCGATCTAAGTTGACCGAT 5690
|||||
QY 1077 CAGTTAGCTCGAATPAGTACAGAAATTTGCTTGACCTTGATGAGAGATCCATGTTTC 1136
|||||
Db 5691 CAGTTAGCTCGAATPAGTACAGAAATTTGCTTGACCTTGATGAGAGATCCATGTTTC 5750
|||||
QY 1137 ATGTTACCTGGAAATGATTTCTATATGCAATTCGAATTCGAATTCGAATTCGAATTCGAAT 1196
|||||
Db 5751 ATGTTACCTGGAAATGATTTGATATGCAATTCGAATTCGAATTCGAATTCGAATTCGAAT 5810
|||||
QY 1197 GAATCTGAACACTGTCTCAATGTTAGATTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255
|||||

Db 5811 GAATCTGAACACTGTCTCAATGTTAGATTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT 5870
|||||
QY 1256 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGACGTTGAACAGAAAGCTAT 1315
|||||
Db 5871 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGACGTTGAACAGAAAGCTAT 5930
|||||
QY 1316 TTTCTGATTCAATCAAGGGTTTATTTGACGTGATTTGAATCTCTTTTGTGTTGTCAGCTCA 1375
|||||
Db 5931 TTTCTGATTCAATCAAGGGTTTATTTGACGTGATTTGAATCTCTTTTGTGTTGTCAGCTCA 5990
|||||
QY 1376 TAAAAA 1381
|||||
Db 5991 TAAAAA 5996
|||||
RESULT 9
ABK89342
ID ABK89342 standard; DNA; 3408 BP.
XX ABK89342;
AC ABK89342;
XX 21-OCT-2002 (first entry)
DT 21-OCT-2002 (first entry)
XX Plasmid pATPCSACT2 DNA.
DE Arsenate reductase; ArsC; cyclic; circular; ds; antimionate; zinc;
KW cadmium; phytochelatin biosynthetic enzyme; arsenate; arsenite; cobalt;
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;
KW industrial waste; phytoremediation.
XX Synthetic.
OS
XX WO200248335-A2.
PN 20-JUN-2002.
XX 13-DEC-2001; 2001WO-US048105.
PD 13-DEC-2000; 2000US-0255001P.
PF 22-JUN-2001; 2001US-0300525P.
PR (UYGE-) UNIV GEORGIA RES FOUND INC.
PA Meagher RB, Li Y;
PI WPI; 2002-583507/62.
XX Recombinant DNA molecules for producing transgenic plants tolerant to heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of contaminated soil or water, encodes arsenate reductase coding sequence.
XX Disclosure; Page 69-71; 131pp; English.
PS The invention relates to a nucleic acid molecule comprising a portion encoding an arsenate reductase coding sequence and a plant-expressible transcription regulatory sequence, the coding sequence being operably linked to the transcription regulatory sequence. The nucleic acid is useful for producing a plant which is resistant to at least one metal ion such as a divalent cadmium ion or antimonate, by introducing the nucleic acid molecule into a plant cell or into plant tissue, selecting for the presence of the nucleic acid molecule to produce a transgenic plant cell or plant tissue and regenerating a plant from the transgenic plant cell or plant tissue. The method further comprises introducing at least one plant expressible phytochelatin biosynthetic enzyme coding sequence into the cell to produce an arsenate metal ion resistant plant. The arsenate reductase coding sequence is expressed under the control of a plant promoter which directs expression in the above ground plant part. The plant is also resistant to cadmium, cobalt, copper, mercury, zinc, antimony, arsenate and arsenite ions. The transgenic plant is useful for bioremediation of arsenate and/or cadmium-contaminated environments, including soil, sediments, mine tailings, water, industrial waste, groundwater and air. The transgenics are also useful for revegetation of

XX The invention relates to a nucleic acid molecule comprising a portion
CC encoding an arsenate reductase coding sequence and a plant-expressible
CC transcription regulatory sequence, the coding sequence being operably
CC linked to the transcription regulatory sequence. The nucleic acid is
CC useful for producing a plant which is resistant to at least one metal ion
CC such as a divalent cadmium ion or antimonate, by introducing the nucleic
CC acid molecule into a plant cell or into plant tissue, selecting for the
CC presence of the nucleic acid molecule to produce a transgenic plant cell
CC or plant tissue and regenerating a plant from the transgenic plant cell
CC or plant tissue. The method further comprises introducing at least one
CC plant expressible phytochelatin biosynthetic enzyme coding sequence into
CC the cell to produce an arsenate resistant plant. The arsenate
CC reductase coding sequence is expressed under the control of a plant
CC promoter which directs expression in the above ground plant part. The
CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc, iron,
CC antimony, arsenate and arsenite ions. The transgenic plant is useful for
CC bioremediation of arsenate and/or cadmium-contaminated environments,
CC including soil, sediments, mine tailings, water, industrial waste,
CC groundwater and air. The transgenics are also useful for revegetation of
CC soils contaminated with metal ions and for removing and sequestering
CC these ions from water, wastewater and aqueous environments. The plants
CC are also useful for phytoremediation of contaminated soil, sediment,
CC water and mine tailings. This sequence represents plasmid pGSAct2 DNA,
CC used in the scope of the invention
XX
SQ Sequence 2857 BP; 804 A; 569 C; 607 G; 877 T; 0 U; 0 Other;

Query Match 89.8%; Score 1297; DB 6; Length 2857;
Best Local Similarity 98.6%; Pred. No. 2.8e-303;
Matches 1371; Conservative 0; Mismatches 12; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAATACATTGATACATATCTCATCTAGATCTAGCTATCATATGTAAG 60
DB |||
QY 9 ATGCTGATCTCAATACATTGATACATATCTCATCTAGATCTAGCTATCATATGTAAG 68
DB |||
QY 61 AAAGTTTTCAGCAATATGNNACGACAAATGCTACACTCGATGTAATGGTATCTCAAC 120
DB |||
QY 69 AAAGTTTTCAGCAATATGCGACGACAAATGCTAGACTCGATGTAATGGTATCTCAAC 128
DB |||
QY 121 TCAACATTATATCTTATACCAACATTAGTAG - CAAAATTTTAAACAACTA - TTTTATGT 178
DB |||
QY 129 TCAACATTATATCTTATACCAACATTAGTAGCAAAATTTTAAACAACTATTTTATGT 188
DB |||
QY 179 ATGCAAGTCTAGCATATGTAATTTGATTCAGAACTCGTTTTCAGCGTTCGGATGTAGT 238
DB |||
QY 189 ATGCAAGTCTAGCATATGTAATTTGATTCAGAACTCGTTTTCAGCGTTCGGATGTAGT 248
DB |||
QY 239 AGTAGCCATTATTTATGTACATCTAATTCGTGAATAGTG-ATATGATGAACATTGTAT 297
DB |||
QY 249 AGTAGCCATTATTTATGTACATCTAATTCGTGAATAGTGATGATGATGAACATTGTAT 308
DB |||
QY 298 CTTATTTGTATAAATPCCATAAACACATCATGAAAGACACTTTCTTCAGGGTCTGAATT 357
DB |||
QY 309 CTTATTTGTATAAATPCCATAAACACATCATGAAAGACACTTTCTTCAGGGTCTGAATT 368
DB |||
QY 358 AATTATGATACAAATCTTATATGAAACGAAATTAATTAATGATTTGATGAATCTTAA 417
DB |||
QY 369 AATTATGATACAAATCTTATATGAAACGAAATTAATTAATGATTTGATGAATCTTAA 428
DB |||
QY 418 TTGAACAGCCAAACACGACGAGGACTAACGTTGCTGGAATGACTCGGTTTAAAGTTAAC 477
DB |||
QY 429 TTGAACAGCCAAACACGACGAGGACTAACGTTGCTGGAATGACTCGGTTTAAAGTTAAC 488
DB |||
QY 478 CACTAAAAAACCGAGCTCTGATGTAACACGCGGATCGAGCGAGTGCATGATGAAGCC 537
DB |||
QY 489 CACTAAAAAACCGAGCTCTGATGTAACACGCGGATCGAGCGAGTGCATGATGAAGCC 548
DB |||
QY 538 ATCAAGCAAAAGACTATATCCAGGGGTGAGATGATTAATTAATTAATTAATTAATTA 597
DB |||
QY 549 ATCAAGCAAAAGACTATATCCAGGGGTGAGATGATTAATTAATTAATTAATTAATTA 608
DB |||
QY 598 CACGAGGGGAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGGTCGAAATGA 656

DB |||
QY 609 CACGAGGGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGGTCGAAATGA 668
DB |||
QY 657 TTCGTGTCTGTGCGATTTTAAATTTTAAATTTTAAAGGCGGAAATTAAGTGTAAAGATA 716
DB |||
QY 669 TTCGTGTCTGTGCGATTTTAAATTTTAAATTTTAAAGGCGGAAATTAAGTGTAAAGATA 728
DB |||
QY 717 AACCCGCTCATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGTCTCTCC 776
DB |||
QY 729 AACCCGCTCATATAAATTCATATATTTTCTCTCCGCTTTTGAATTTGTCTCGTTGTCTCTCC 788
DB |||
QY 777 TCATTTTCATCAGCGGTTTGAATCTCCGGGCACTTTGACAGAGAAACAAGAAAGATA 836
DB |||
QY 789 TCATTTTCATCAGCGGTTTGAATCTCCGGGCACTTTGACAGAGAAACAAGAAAGATA 848
DB |||
QY 837 CTAAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTTTCGAATCTTCTCTCA 896
DB |||
QY 849 CTAAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTTTCGAATCTTCTCTCA 908
DB |||
QY 897 TCTCATCTTCTTCTCCGCTCTTTTCTTCCAGGTAATAGGAACCTTTCTGGAATCTACTTT 956
DB |||
QY 909 TCTCA---TCTTCTTCCGCTCTTTTCTTCCAGGTAATAGGAACCTTTCTGGAATCTACTTT 965
DB |||
QY 957 ATTGCTCGATCTCGATCTTGTCTTCAATTTCTTCTGAGATCTGGAATCTGTTTAAATTT 1016
DB |||
QY 966 ATTGCTCGATCTCGATCTTGTCTTCAATTTCTTCTGAGATCTGGAATCTGTTTAAATTT 1025
DB |||
QY 1017 GGATCTGTGAACCTCCACTTAATCTTTTGGTTTCTTACTAGATCCGATCTAAGTTGACCGAT 1076
DB |||
QY 1026 GGATCTGTGAACCTCCACTTAATCTTTTGGTTTCTTACTAGATCCGATCTAAGTTGACCGAT 1085
DB |||
QY 1077 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTGGACTTGTAGAGAGATCCATGTTTC 1136
DB |||
QY 1086 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTGGACTTGTAGAGAGATCCATGTTTC 1145
DB |||
QY 1137 ATGTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACTGTTGAAAGTTAGATT 1196
DB |||
QY 1146 ATGTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACTGTTGAAAGTTAGATT 1205
DB |||
QY 1197 GAATCTGACACTGCAATGTTAGATTTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255
DB |||
QY 1206 GAATCTGACACTGCTCAATGTTAGATTTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT 1265
DB |||
QY 1256 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTTCGTACGTTGAACAGAAAGCTAT 1315
DB |||
QY 1266 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTTCGTACGTTGAACAGAAAGCTAT 1325
DB |||
QY 1316 TTCTGATTCATTCAGGGTTTATTTGACTGATTTGAATCTTTTTTGTGTGTTTGCAGCTCA 1375
DB |||
QY 1326 TTCTGATTCATTCAGGGTTTATTTGACTGATTTGAATCTTTTTTGTGTGTTTGCAGCTCA 1385
DB |||
QY 1376 TAAAAAATGGC 1386
DB |||
QY 1386 TAAACCATGGC 1396
DB |||

RESULT 11
AAD09791
ID AAD09791 standard; DNA; 1219 BP.
XX
AC AAD09791;
XX
DT 10-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana Act2 promoter including intron of Act2 gene.
XX
KW DNA construct; structural gene; SG; non-translated region; NTR; actin;
KW Act; elongation factor 1 alpha; EFlalpha; weed control; BPSP synthase;
KW glyphosate tolerance; glyphosate oxidoreductase; herbicide tolerance;
KW PCR primer; Act2 promoter; ds.
XX
OS Arabidopsis thaliana.
XX

10-SEP-2001 (first entry)
Chimeric CamV-Act2 promoter including first intron of Act2 gene.
DNA construct; structural gene; SG; non-translated region; NTR; actin;
Act; elongation factor 1 alpha; EFlalpha; weed control; EPSP synthase;
glyphosate tolerance; glyphosate oxidoreductase; herbicide tolerance;
PCR primer; Act2 promoter; Cauliflower mosaic virus; CamV; chimeric; ds.
Cauliflower mosaic virus.
OS Arabidopsis thaliana.
OS Chimeric.
FH Key Location/Qualifiers
FT promoter 1..523 /tag= a
FT /note= "CamV promoter"
FT misc_feature 534..1742
FT /tag= b
FT /note= "Arabidopsis thaliana Act2 promoter, intron and 5'
FT UTR region"
XX
PN WO200144457-A2.
XX
XX 21-JUN-2001.
XX
XX 12-DEC-2000; 2000WO-US033633.
XX
XX 16-DEC-1999; 99US-0171173P.
XX (MONS) MONSANTO CO.
XX Fincher KL, Flasiniski S, Wilkinson JQ;
PI
XX WPI; 2001-408480/43.
XX
XX Novel DNA construct for controlling weeds, containing Arabidopsis actin
promoter sequences and elongation factor 1-alpha promoter sequences
operably linked to structural genes that function in crop plants.
XX
XX Claim 14; Page 100-101; 101pp; English.
XX
XX The present invention relates to DNA construct comprising an expression
cassette containing a promoter DNA sequence operably linked to a
structural gene (SG) encoding an agronomically useful protein, and a 3'
non-translated region (NTR) that functions in plants to cause the
addition of polyadenylated nucleotides to the 3' end of the RNA sequence,
operably linked to SG. The present invention further relates to plant
expression constructs that comprise Arabidopsis actin (Act) promoter
sequences Act1a, Act1b, Act2, Act3, Act7, Act8, Act11, Act12, elongation
factor 1 alpha (EFlalpha) promoter sequence, fragments and cis elements
derived from these promoters operably linked to heterologous structural
gene sequences that function in crop plant cells. The DNA constructs are
useful for controlling weeds, by providing a crop plant transformed with
a DNA construct comprising promoter DNA sequence that is functional in a
plant cell, a glyphosate tolerance gene (which is an EPSP synthase gene
or a glyphosate oxidoreductase gene) or a herbicide tolerance gene. The
present sequence is chimeric Cauliflower mosaic virus (CamV)-Act2
promoter including first intron of Arabidopsis thaliana Act2 gene.
(Updated on 11-SEP-2003 to standardise OS field)
XX Sequence 1742 BP; 527 A; 327 C; 362 G; 526 T; 0 U; 0 Other;
Query Match 81.3%; Score 1174; DB 4; Length 1742;
Best Local Similarity 99.3%; Pred. No. 1.3e-273;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
XX 164 CAACATATTTTATGATGCAAGAGTCAGCATATGATATTCATTCAGAAATCGTTTGAC 223
DB 533 CAACATATTTTATGATGCAAGAGTCAGCATATGATATTCATTCAGAAATCGTTTGAC 592
XX 224 GAGTTCCGATGTAGTAGCCATTTATTTAAATGTACATCAATCGTGAATAGTATG 283
|||||

Db 593 GAGTTCCGATGTAGTAGCCATTTATTTAAATGTACATCAATCGTGAATAGTATG 652
Qy 284 ATGAACAATTCATCTTATTTGTAATAATATCATATAACACATCATGAAGACACTTTCTT 343
Db 653 ATGAACAATTCATCTTATTTGTAATAATATCATATAACACATCATGAAGACACTTTCTT 712
Qy 344 TCAGGCTCTGAATTAATTTATGATACAAATCTTAATAGAAAAAGAAATTAATTAATTCGTTGAAT 403
Db 713 TCAGGCTCTGAATTAATTTATGATACAAATCTTAATAGAAAAAGAAATTAATTAATTCGTTGAAT 772
Qy 404 TGTATGAAATCTAATTTGAAGCAAGCAACGACGAGGAGCTAAACGTTCCGCTGGATGACT 463
Db 773 TGTATGAAATCTAATTTGAAGCAAGCAACGACGAGGAGCTAAACGTTCCGCTGGATGACT 832
Qy 464 CGTTTAAAGTTAAACCACTAAATAAAGCGAGCTGTCATGTAAACGCGATCGAGCAGGTC 523
Db 833 CGTTTAAAGTTAAACCACTAAATAAAGCGAGCTGTCATGTAAACGCGATCGAGCAGGTC 892
Qy 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAAATCAAGGGGTGAGATGATTAATTAATAGTT 583
Db 893 ACAGTCATGAAGCCATCAAGCAAAAGAACTAAATCAAGGGGTGAGATGATTAATTAATAGTT 952
Qy 584 TAAAAATTTAGTTAAACGAGGAGAAA-GCTGTCTGACAGCAGGTCACGTTATCTTTTACC 642
Db 953 TAAAAATTTAGTTAAACGAGGAGAAAAGGCTGTCTGACAGCAGGTCACGTTATCTTTTACC 1012
Qy 643 TGTGTCGAAATGATTCGTGTCGATTTTAATTTATTTTGAAGGCCGAAAAATAA 702
Db 1013 TGTGTCGAAATGATTCGTGTCGATTTTAATTTATTTTGAAGGCCGAAAAATAA 1072
Qy 703 AGTTCTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTG 762
Db 1073 AGTTCTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCTCCGCTTTGAAATTTG 1132
Qy 763 TCTCGTTGTCCTCTCACTTTTCATCAGCGGTTTTGAAATCTCGGCGACTTGACAGAGAAG 822
Db 1133 TCTCGTTGTCCTCTCTCACTTTTCATCAGCGGTTTTGAAATCTCGGCGACTTGACAGAGAAG 1192
Qy 823 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT 882
Db 1193 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT 1252
Qy 883 TGAATCTTCTCAATCTCATCTTCTTCTCCGCTCTTTTCTTCCAGGTAATAGGAACCTT 942
Db 1253 TGAATCTTCTCAATCTCA---TCTTCTCCGCTCTTTTCTTCCAGGTAATAGGAACCTT 1309
Qy 943 TCTGATCTACTTTATTTGCTGGATCTCGATCTGTTTTTCTCAATTTCTTGAGATCTGG 1002
Db 1310 TCTGATCTACTTTATTTGCTGGATCTCGATCTGTTTTTCTCAATTTCTTGAGATCTGG 1369
Qy 1003 AATTGTTTTTATTTGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAAATCGAT 1062
Db 1370 AATTGTTTTTATTTGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAAATCGAT 1429
Qy 1063 CTAAGTTGACCCGATCAGTTAGCTCGATTTATAGTACCAGAAATTTGGCTTTGACCTTTGATGG 1122
Db 1430 CTAAGTTGACCCGATCAGTTAGCTCGATTTATAGTACCAGAAATTTGGCTTTGACCTTTGATGG 1489
Qy 1123 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCT 1182
Db 1490 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCT 1549
Qy 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1242
Db 1550 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1609
Qy 1243 GTTATGATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGAATTTGTAGTGTCTGATGTTG 1302
Db 1610 GTTATGATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGAATTTGTAGTGTCTGATGTTG 1669
Qy 1303 AACAGAAAGCTATTTCTGATTTCAATCAGGTTTATTTGATGATTTGAACTCTTTTGTG 1362
Db 1670 AACAGAAAGCTATTTCTGATTTCAATCAGGTTTATTTGATGATTTGAACTCTTTTGTG 1729

QY 1183 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1242
DB |||||||
QY 1243 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1017
DB |||||||
QY 1243 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1302
DB |||||||
QY 1018 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1077
DB |||||||
QY 1303 AACAGAAAGCTATTTCTGATTAATCAATCAGGGTTTATTTGACTGTATTTGAATCTTTTGTG 1362
DB |||||||
QY 1078 AACAGAAAGCTATTTCTGATTAATCAATCAGGGTTTATTTGACTGTATTTGAATCTTTTGTG 1137
DB |||||||
QY 1363 TGTTCACGTCATATAAAATGGCTCAGGCTGA 1395
DB |||||||
QY 1138 TGTTCACGTCAGGATCCATCGATAGCTTTA 1170
DB |||||||

RESULT 14

ABK52080
ID ABK52080 standard; DNA; 1259 BP.
XX
AC ABK52080;
XX
DT 13-AUG-2002 (first entry)
XX
DE Modified plant promoter #5.
XX
KW Promoter; ds; plant; soybean; peanut; alfalfa; tomato; eggplant; potato;
KW cabbage; turnip; rapeseed; apple; pear; berry; cucumber; carrots;
KW transgenic plant.
XX
OS Synthetic.
XX
PN US2002049992-A1.
XX
PD 25-APR-2002.
XX
PF 22-JUN-2001; 2001US-00887384.
XX
PR 01-DEC-1999; 99US-00453366.
XX
PA (HAMI/) HAMILTON C.
XX
PI Hamilton C;
XX
XX WPI; 2002-434778/46.
DR
PT Modified plant promoter DNA sequences, useful for expressing genes in
PT monocot and dicotyledon plants, e.g. soybean, tomato, cabbage, apples,
PT cucumbers, beets and carrots.
XX

Claim 1; Page 11; 18pp; English.

This invention relates to novel modified plant promoters given in the
CC specification, their complements or double stranded equivalents. The
CC modified DNA promoters can be efficiently expressed in higher eukaryotes
CC (e.g. plants), and used to direct gene expression within them. In
CC particular the modified promoters are efficiently expressed in
CC dicotyledonous plants, e.g. species of legumes from the family Fabaceae
CC (such as soybean, peanut, and alfalfa), species of the Solanaceae family
CC (e.g. tomato, eggplant and potato), species of the family Brassicaceae
CC (e.g. cabbage, turnips and rapeseed), species of the family Rosaceae
CC (e.g. apples, pears and berries), and members of the families
CC Cucurbitaceae (e.g. cucumbers), Chenopodiaceae (e.g. beets) and
CC Umbelliferae (e.g. carrots). Although effective, the promoters have not
CC been modified or optimized to provide enhanced or improved
CC characteristics or traits. The DNA plant promoters have been modified to
CC advantageously provide improved characteristics or traits in plants and
CC are efficiently expressed in higher eukaryotes (e.g. plants). The DNA
CC constructs allow for the preparation of stably transformed cells
CC expressing heterologous proteins, and for the subsequent regeneration of
CC fertile, transgenic plants and progeny containing desired modified
CC promoters. The present sequence represents a modified plant promoter of

CC the invention

XX
SQ Sequence 1259 BP; 380 A; 204 C; 246 G; 429 T; 0 U; 0 Other;
Query Match 76.5%; Score 1104.6; DB 6; Length 1259;
Best Local Similarity 94.0%; Pred. No. 7.3e-257;
Matches 1198; Conservative 0; Mismatches 59; Indels 18; Gaps 4;
QY 164 CAACCTATTTTATGATGTCAGAGGTCAGCATATGATTAATGATTCAGAAATCGTTTTCAG 223
DB 1 CAACCTATTTTATGATGTCAGAGGTTTCATCTTGATTAATGATTCAGAAATCGTTTTCAG 60
QY 224 GAGTTCCGATGTAGTAGTCCATTTATTAATGTATACATCACTAATCGTGAATAGTATG 283
DB 61 GAGTTCCGATGTAGTAGTCCATTTATTAATGTATACATCACTAATCGTGAATAGTATG 120
QY 284 ATGAAACATTTGATCTTTATTTGATATAATATCATATAACATCATGATGAAAGACATTTCTT 343
DB 121 ATGAAACATTTGATCTTTATTTGATATAATATCATATAACATCATGATGAAAGACATTTCTT 180
QY 344 TCAGGGTCTGATTAATTAATGATACAATTTCTTAATAGAAAACGAATTAATTAACGTTTGAAT 403
DB 181 TCACGGTCTGAATTAATTAATGATACAATTTCTTAATAGAAAACGAATTAATTAACGTTTGAAT 240
QY 404 TGTATGAAATCTAATTTGAACCAAGCCAAACGACGAGGACTAAACGTTGCTCGATTGACT 463
DB 241 TGTATGAAATCTAATTTGAACCAAGCCAAACGACGAGGACTAAACGTTGCTCGATTGACT 300
QY 464 CGTTTAAAGTTAAACCACTAAACCAAGCGAGGCTCATGTAAACACGCGGATCGAGGTC 523
DB 301 AAGTTTAAAGATAAACCACTAAACCAAGCGAGGCTCATGTAAACACGCGGATCGAGGTC 360
QY 524 ACAGTCATGAGCCATCAACGACCAAGCAACTAATCCAGGGGTGAGATTAATTAATGTT 583
DB 361 ACAGTCATGAGCCATCAACGACCAAGCAACTAATCCAGGGGTGAGATTAATTAATGTT 420
QY 584 TAAAAATTTAGTTTAAACACGAGGGGAAA-GCTGTCTGCACAGCCAGGTCACGTTATCTTACC 642
DB 421 TAAAAATTTAGTTTAAACACGAGGGGAAAAGGCTGTCTGCACAGCCAGGTCACGTTATCTTACC 480
QY 643 TGTGTGCAATGATTCGTGTCTGCAATTTTAAATTTATTTTGAAGGCCGAAAATTA 702
DB 481 TGTGTGCAATGATTCGTGTCTGCAATTTTAAATTTATTTTGAAGGCCGAAAATTA 540
QY 703 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATG 762
DB 541 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATG 600
QY 763 TCTCGTTGTCTCTCCTCACCTTTTCATCAGCCGTTTTCGAAATCTCCGCGGACTTGCACAGAGAAG 822
DB 601 TCTCGTTGTCTCTCCTCACCTTTTCATCAGCCGTTTTCGAAATCTCCGCGGACTTGCACAGAGAAG 660
QY 823 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTT 882
DB 661 AACAGGAAG-----AGAGAGAAAGTAAGAGATAATCCAGG-----TTCTCGTTT 706
QY 883 TGAATCTTCTCAATCTCATCTTCTTCTTCTCCGCTCTTCTTCTTCCAGGTAATAGAACTT 942
DB 707 TGAATCTTCTCAATCTCA-----TCTTCTTCTCCGCTCTTCTTCTTCCAGGTAATAGAACTT 763
QY 943 TCTCGATCTACCTTTATTTTGTGTCGATCTCGATCTTCTTCTTCTCAATTTCTCGATCTG 1002
DB 764 TCTCGATCTACCTTTATTTTGTGTCGATCTCGATCTTCTTCTTCTCAATTTCTCGATCTG 823
QY 1003 AATCTGTTTAAATTTGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTACTAGAACTCGAT 1062
DB 824 TTTTCTGTTTAAATTTGATGTAATTTAGATCACTAAATCTTTTGGTTTCTACTAGAACTCGAT 883
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATATAGTACCAGAAATTTGCTTGTGACCTTGATG 1122
DB 884 CTAAGTTGACCGATCAGTTAGCTCGATATAGTACCAGAAATTTGCTTGTGACCTTGATG 943
QY 1123 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTTGTATGTGAATTTGAAATCTGAAT 1182

Db 944 AGAGATCCATGTTTCATGTTACCTGGGAATGATTTGTATATGGAATGAAATCTGAAT 1003
QY 1183 GTTGAAGTTAGATGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1242
Db 1004 GTTGAAGTTAGATTAATGACTGTAACTGTCAATGTTAGATTAATGACTGTAACTGTTTAA 1063
QY 1243 GTTAGATGAAGTTTGTGTAGATTTCTCGAACTTGTAGATTTGTAGTGTGCTAGCTTTG 1302
Db 1064 GTTAGATGAAGTTTGTGTAGATTTCTCGAACTTGTAGATTTGTAGTGTGCTAGCTTTG 1123
QY 1303 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATTTGACTCTTTTGTG 1362
Db 1124 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATTTGACTCTTTTGTG 1183
QY 1363 TGTTCGAGCTCATRAAAAATGCTGAGGCTGACGATATTTCAACCAATCGTGTGTGACAA 1422
Db 1184 TGTTCGAGCTCATRAAACCATGCTGAGGCTGATGATATTTCAACCAATCGTGTGTGACAA 1243
QY 1423 TGGTACTGGAATGGT 1437
Db 1244 TGGTACTGGAATGGT 1258

RESULT 15
AAD01288
ID AAD01288 standard; DNA; 15676 BP.
XX
AC AAD01288;
XX
DT 12-OCT-2000 (first entry)
XX
DE Arabidopsis transformation binary vector, pAct2-bin.
XX
KW Matrix Attachment Region; MAR; scaffold attachment region;
KW Arabidopsis transformation binary vector; pAct2-bin; gene expression;
KW transgenic organism; ds.
XX
OS Synthetic.
XX
PN WO200032800-A1.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US028123.
XX
PR 01-DEC-1998; 98US-0110437P.
XX
PA (DOWC) DOW AGROSCIENCES LLC.
XX
PI Van Der Geest AHM, Ainley WM, Cowen NM, Welter ME, Woosley AT;
XX
DR WPI; 2000-412345/35.
XX
PT An isolated DNA molecule for use as a matrix attachment region to
PT increase expression of genes introduced in transformed plants comprises a
PT 298 base pair sequence described in the specification.
XX
PS Example 3; Page 55-59; 73pp; English.

CC The patent discloses a DNA molecule, useful as matrix attachment region
CC (MAR) or scaffold attachment region, to increase the expression of genes
CC introduced in transformed plants. MARs are located in non-transcribed
CC regions of genes and form the physical boundaries of individual DNA
CC loops. They are rich in adenosine and thymine bases and contain certain
CC conserved sequence elements and structural features. They are about 300-
CC 2000 bp in length. Increased levels of expression of DNA introduced into
CC plants can be achieved by use of MAR. They can also reduce the position
CC effect in transgenic organisms. The present DNA sequence is the
CC Arabidopsis transformation binary vector pAct2-bin, that contains a
CC Act2/GUS/nos cassette (Act2 transcription initiation region/GUS
CC structural gene/nos 3' untranslated region), 19S/NPTII/orf25polyA as a
CC selectable marker and 35S/GFP/nos as an independent reporter gene. This

CC binary vector is used to construct other vectors, used to test the two
CC orientations of the artificial MAR dimers in Arabidopsis
XX
SQ Sequence 15676 BP; 4226 A; 3741 C; 3838 G; 3868 T; 0 U; 3 Other;
Query Match 74.7%; Score 1078.8; DB 3; Length 15676;
Best Local Similarity 98.3%; Pred. No. 2.7e-250;
Matches 1154; Conservative 0; Mismatches 12; Indels 8; Gaps 6;
QY 205 GATTCAGAAATCTTTTTCAGGAGTTCGGATGTAGTAGTCCCAATTTTAAATGACATCT 264
Db 84 GAATTCGGCTTGTGTTGACGAGTTCGGATGTAGTAGTCCCAATTTTAAATGACATCT 143
QY 265 AATCGTGAATAGTG-ATATGATGAACAATTTGATCTTATTTGATATAAATCCATAAACAC 323
Db 144 AATCGTGAATAGTGATGAATGATGAACAATTTGATCTTATTTGATATAAATCCATAAACAC 203
QY 324 ATCATGAAGAAGACACATTTCTTTCAGGCTCGAATTAATTAATGATACAAATTTCTAATAGAAA 383
Db 204 ATCATGAAGAAGACATTTCTTTCAGGCTCGAATTAATTAATGATACAAATTTCTAATAGAAA 263
QY 384 CGAATTAATTAATGATGATGAATTTGATGAATCTAAATTTGAACAGCCACACACGACGAC 443
Db 264 CGAATTAATTAATGATGATGAATTTGATGAATCTAAATTTGAACAGCCACACACGACGAC 323
QY 444 TAACGTTGCTGGATTTGACTCGGTTTAAAGTTAAACCACTAAATAAAGCGAGCTGTCTATGTA 503
Db 324 TAACGTTGCTGGATTTGACTCGGTTTAAAGTTAAACCACTAAATAAAGCGAGCTGTCTATGTA 383
QY 504 ACACGCGGATCGAGCAGGTCTACAGTCATGAGCCATCAAGCAAAAGAACTAATCCAAGG 563
Db 384 ACACGCGGATCGAGCAGGTCTACAGTCATGAGCCATCAAGCAAAAGAACTAATCCAAGG 443
QY 564 GG-TGAGATGATTAATTAATTTAAATAATTTAGTTTAAACAGCGGGGAAA-GCTCTCTGACAG 621
Db 444 GGCTGAGATGATTAATTAATTTAAATAATTTAGTTTAAACAGCGGGGAAAAGCTCCCTGACAG 503
QY 622 CCAGTCACTGTTATCTTTTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
Db 504 CCAGTCACTGTTATCTTTTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563
QY 682 TTTTGAAGGCGGAAATAAAGTTGTAAGAGATAAACCCTGCTATATAAATTCATATAT 741
Db 564 TTTTGAAGGCGGAAATAAAGTTGTAAGAGATAAACCCTGCTATATAAATTCATATAT 623
QY 742 TTTCTCTCCGCTTTGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
Db 624 TTTCTCTCCGCTTTGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 683
QY 802 TCCGCGGACTTTGACAGAGAGAAACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 861
Db 684 TCCGCGGACTTTGACAGAT 743
QY 862 CAGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCATCTTCTTCTTCTCCGCTCTTTC 921
Db 744 CAGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCA---TCTTCTCCGCTCTTTC 800
QY 922 TTTCCAGGTAATAGGAACTTTCTGGATCTACTTTTATTTGCTGGATCTCGATCTTGTGTTT 981
Db 801 TTTCCAGGTAATAGGAACTTTCTGGATCTACTTTTATTTGCTGGATCTCGATCTTGTGTTT 860
QY 982 CTCAATTTCTTTGAGATCTGGAAATTCGTTTAAATTT--GGATCTGTGAACCTCCACATAAATC 1040
Db 861 CTCAATTTCTTTGAGATCTGGAAATTCGTTTAAATTTGGGATCTGTGAACCTCCACATAAATC 920
QY 1041 TTTTGGTTTAACTAGATTCGATCTAAGTTGACCGATCAGTTAGCTCGATATAGCTACCA 1100
Db 921 TTTTGGTTTAACTAGATTCGATCTAAGTTGACCGATCAGTTAGCTCGATATAGCTACCA 980
QY 1101 GAATTTGCT 1160
Db 981 GAATTTGCT 1040

```
QY 1161 TATGTGAATTGAAATCTGAACCTGTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAG 1220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1041 TATGTGAATGAAATCTGAACCTGTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAG 1100
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1221 ATTGAATCTGAACACTGTGTTTAA-GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTT 1279
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1101 ATTGAATCTGAACACTGTGTTTAAAGTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTT 1160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1280 AGGATTTGTAGTGTCTGACGTTGAACAGAAAGCTATTTCTGATTCAATCAGGGTTTATTT 1339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1161 AGGATTTGTAGTGTCTGACGTTGAACAGAAAGCTATTTCTGATTCAATCAGGGTTTATTT 1220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1340 GACTGTATTGAACTCTTTTGTGTGTTTGCAGCT 1373
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1221 GACTGTATTGAACTCTTTTGTGTGTTTGCAGCT 1254
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: May 27, 2005, 20:09:22
Job time : 855 secs

1	1174.8	81.4	1219	3	US-09-737-6988-9	Sequence 9, Appli
2	1174.8	81.4	1219	4	US-09-737-628A-9	Sequence 9, Appli
3	1174	81.3	1742	3	US-09-737-6988-30	Sequence 30, Appl
4	1174	81.3	1742	4	US-09-737-626A-30	Sequence 30, Appl
5	75.4	5.2	1271	3	US-09-737-6988-10	Sequence 10, Appl
6	75.4	5.2	1271	4	US-09-737-626A-10	Sequence 10, Appl
7	73.8	5.1	1800	3	US-09-737-6988-29	Sequence 29, Appl
8	73.8	5.1	1800	4	US-09-737-626A-29	Sequence 29, Appl
9	52.8	3.7	1141	4	US-09-806-7088-22	Sequence 22, Appl
10	49	3.4	19438	1	US-09-949-016-13699	Sequence 12699, A
11	48.6	3.4	731	1	US-08-451-405A-2	Sequence 2, Appli
12	48.2	3.3	95255	4	US-09-949-016-17067	Sequence 17067, A
13	47.8	3.3	5203	3	US-09-257-770-1	Sequence 1, Appli
14	47.2	3.3	109038	4	US-09-949-016-12199	Sequence 12199, A
15	47	3.3	1141	4	US-09-806-7088-22	Sequence 22, Appl
16	46.6	3.2	18773	4	US-09-949-016-14164	Sequence 14164, A
17	46.6	3.2	640681	4	US-09-790-988-1	Sequence 1, Appli
18	46.2	3.2	2448	1	US-08-526-964-2	Sequence 2, Appli
19	46.2	3.2	2448	2	US-08-946-617-2	Sequence 2, Appli
20	46.2	3.2	2448	3	US-09-031-897-2	Sequence 2, Appli
21	46.2	3.2	30820	4	US-09-949-016-17145	Sequence 17145, A
22	46	3.2	134987	4	US-09-949-016-15348	Sequence 15348, A
23	46	3.2	134987	4	US-09-949-016-15349	Sequence 15349, A
24	46	3.2	134987	4	US-09-949-016-15350	Sequence 15350, A
25	46	3.2	134987	4	US-09-949-016-15507	Sequence 15507, A
26	46	3.2	134987	4	US-09-949-016-15508	Sequence 15508, A
27	46	3.2	134987	4	US-09-949-016-15509	Sequence 15509, A

```
Db 301 |||||TAAGTTAACTAACCACTAAAAAACCGAGCTGTCATGTAAACACGCGATCAGCAGGTC 360
Qy 524 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCNAGGGGTGAGATTAATAGTT 583
Db 361 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCNAGGGGTGAGATTAATAGTT 420
Qy 584 TAAAAATTAGTTAAACACGAGGGAATAA-GCTGTCACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAGTTAAACACGAGGGAATAAAGGCTGTCACAGCCAGGTCACGTTATCTTTACC 480
Qy 643 TGTGTCGAAATGATTCGTGTCGTGATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 702
Db 481 TGTGTCGAAATGATTCGTGTCGTGATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 540
Qy 703 AGTTGTAAGAGATAAACCCTCATATAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 762
Db 541 AGTTGTAAGAGATAAACCCTCATATAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 600
Qy 763 TCTCGTTGTCCTCCCTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACT 822
Db 601 TCTCGTTGTCCTCCCTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACT 660
Qy 823 AACAAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT 882
Db 661 AACAAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT 720
Qy 883 TGAATCTCTCAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
Db 721 TGAATCTCTCAATCTCA---TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777
Qy 943 TCTGATCTACTTTTATTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1002
Db 778 TCTGATCTACTTTTATTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 837
Qy 1003 AATTCTGTTTAAATTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1062
Db 838 AATTCTGTTTAAATTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 897
Qy 1063 CTAAGTTGACCATGATTTAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1122
Db 898 CTAAGTTGACCATGATTTAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 957
Qy 1123 AGAGATCCATGTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTT 1182
Db 958 AGAGATCCATGTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTT 1017
Qy 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCATGTTAGATTTGAATCTGAACACTGTTTAA 1242
Db 1018 GTTGAAGTTAGATTTGAATCTGAACACTGTCATGTTAGATTTGAATCTGAACACTGTTTAA 1077
Qy 1243 GTTAGATGAAGTTTGTGATAGATTTCTCGAACCCTTAGATTTGTAGTGTGATGTTG 1302
Db 1078 GTTAGATGAAGTTTGTGATAGATTTCTCGAACCCTTAGATTTGTAGTGTGATGTTG 1137
Qy 1303 AACAGAACTATTTCTGATTTCAATCAGGTTTATTTGATCTGATTTGATCTGATTTGATCTGATTTG 1362
Db 1138 AACAGAACTATTTCTGATTTCAATCAGGTTTATTTGATCTGATTTGATCTGATTTGATCTGATTTG 1197
Qy 1363 TGTGTTGACGTCATAAAAAATG 1384
Db 1198 TGTGTTGACGACACTCACCATG 1219
```

RESULT 2

```
US-09-737-626A-9
; Sequence 9, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiński, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911e1 Plant Expression Constructs
```

```
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-09-737-626A-9
```

```
Query Match 81.4%; Score 1174.8; DB 4; Length 1219;
Best Local Similarity 98.7%; Pred. No. 6.7e-309;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
```

```
Qy 164 CAACATATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 223
Db 1 CAACATATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy 224 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
Db 61 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Qy 284 ATGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343
Db 121 ATGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 344 TCAGGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
Db 181 TCAGGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 404 TGTATGAAATCTAAATGAAACCAAGCCAAACGACGACGACGACGACGACGACGACGACGACGACGACG 463
Db 241 TGTATGAAATCTAAATGAAACCAAGCCAAACGACGACGACGACGACGACGACGACGACGACGACGAC 300
Qy 464 CGTTTAAAGTTAAACCACTAAACCAAGCCAAACGACGACGACGACGACGACGACGACGACGACGACG 523
Db 301 CGTTTAAAGTTAAACCACTAAACCAAGCCAAACGACGACGACGACGACGACGACGACGACGACGACG 360
Qy 524 ACAGTCATGACCATCAAGCAAGAAAGAACTAATCCNAGGGGTGAGATTAATAGTT 583
Db 361 ACAGTCATGACCATCAAGCAAGAAAGAACTAATCCNAGGGGTGAGATTAATAGTT 420
Qy 584 TAAAAATTAGTTAAACACGAGGGAATAA-GCTGTCACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAGTTAAACACGAGGGAATAAAGGCTGTCACAGCCAGGTCACGTTATCTTTACC 480
Qy 643 TGTGTCGAAATGATTCGTGTCGTGATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 702
Db 481 TGTGTCGAAATGATTCGTGTCGTGATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 540
Qy 703 AGTTGTAAGAGATAAACCCTCATATAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 762
Db 541 AGTTGTAAGAGATAAACCCTCATATAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 600
Qy 763 TCTCGTTGTCCTCCCTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACT 822
Db 601 TCTCGTTGTCCTCCCTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACT 660
Qy 823 AACAAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT 882
Db 661 AACAAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT 720
Qy 883 TGAATCTCTCAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
Db 721 TGAATCTCTCAATCTCA---TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777
```

QY 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTGAGATCTGG 1002
DB 778 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTGAGATCTGG 837
QY 1003 AATTGCTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGATCGAT 1062
DB 838 AATTGCTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGATCGAT 897
QY 1063 CTAAGTTGACCGATCAGTTAGTTCGATTATAGCTACCCAGAAATTTGGCTTGACCTTGATGG 1122
DB 898 CTAAGTTGACCGATCAGTTAGTTCGATTATAGCTACCCAGAAATTTGGCTTGACCTTGATGG 957
QY 1123 AGAGATCCATGTTTCATGTTACCTGGGAATGATTTGTATGTGATGTAATGAAATCTGAAC 1182
DB 958 AGAGATCCATGTTTCATGTTACCTGGGAATGATTTGTATGTGATGTAATGAAATCTGAAC 1017
QY 1183 GTTGAAGTTAGATTTGAATCTGAACCTGTCAATGTTAGATTTCAATCTGAACACTGTTTAA 1242
DB 1018 GTTGAAGTTAGATTTGAATCTGAACCTGTCAATGTTAGATTTCAATCTGAACACTGTTTAA 1077
QY 1243 GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGCTTG 1302
DB 1078 GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGCTTG 1137
QY 1303 AACAGAAAGCTATTCTGATTTCAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTG 1362
DB 1138 AACAGAAAGCTATTCTGATTTCAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTG 1197
QY 1363 TGTGTCAGCTCATAAATAATG 1384
DB 1198 TGTGTCAGCTCATCAACATG 1219

RESULT 3

US-09-737-698B-30
; Sequence 30, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Wilkinsons, Jack
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737,698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 polynucleotides + Act2 in
; OTHER INFORMATION: tro
US-09-737-698B-30

Query Match 81.3%; Score 1174; DB 3; Length 1742;
Best Local Similarity 99.3%; Pred. No. 1.3e-308;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGATTAATTCATTCAGAATCGTTTTGAC 223
DB 533 CAACATATTTTATGATGCAAGAGTCAGCATATGATTAATTCAGAATCGTTTTGAC 592
QY 224 GAGTTCGGATGTAGTAGTACCATTTATTAATGTACATACATAATCGTGAATAGTATG 283
DB 593 GAGTTCGGATGTAGTAGTACCATTTATTAATGTACATACATAATCGTGAATAGTATG 652
QY 284 ATGAACAATGTATCTTATTTGTATTAATATCATTAACACATCATGAAGACACTTTCTT 343

DB 653 ATGAACAATGTATCTTATTTGTATTAATATCATTAACACATCATGAAGACACTTTCTT 712
QY 344 TCAGGGTCTGAATTAATTTATGATACAAATCTAATAGAAAAACGAAATTAATTAACGTTGAAT 403
DB 713 TCAGGGTCTGAATTAATTTATGATACAAATCTAATAGAAAAACGAAATTAATTAACGTTGAAT 772
QY 404 TGTATGAAATCTAATTTGAACCAAGCCAAACGACGAGGACTAAGCGTTGCTCGATGACT 463
DB 773 TGTATGAAATCTAATTTGAACCAAGCCAAACGACGAGGACTAAGCGTTGCTCGATGACT 832
QY 464 CGGTTTAAAGTTAAACCACTAATAAACAACGAGCTGTCATGTAAACACGCGGATCCAGCAGGTC 523
DB 833 CGGTTTAAAGTTAAACCACTAATAAACAACGAGCTGTCATGTAAACACGCGGATCCAGCAGGTC 892
QY 524 ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTTAGTT 583
DB 893 ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTTAGTT 952
QY 584 TAAAAATTTAGTTAAACACGAGGAAAAA--GCTGTCTGACAGCCAGGTCAGCTTATCTTTACC 642
DB 953 TAAAAATTTAGTTAAACACGAGGAAAAAGGCTGTCTGACAGCCAGGTCAGCTTATCTTTACC 1012
QY 643 TGTGTCTCGAAATGATTTCTGTCTGTCTGATTTTAAATTTTAAAGGCGCGAAAAATAA 702
DB 1013 TGTGTCTCGAAATGATTTCTGTCTGTCTGATTTTAAATTTTAAAGGCGCGAAAAATAA 1072
QY 703 AGTTCTAAGAGATAAACCCGCTATATATAATTTATATATTTTCTCCCGCTTTGAAATTTG 762
DB 1073 AGTTCTAAGAGATAAACCCGCTATATATAATTTATATATTTTCTCCCGCTTTGAAATTTG 1132
QY 763 TCTCGTTGCTCTCTCACTTTTCACTGCGGCTTTTGAATCTCTCGGCGACTTGAACAGAG 822
DB 1133 TCTCGTTGCTCTCTCTCACTTTTCACTGCGGCTTTTGAATCTCTCGGCGACTTGAACAGAG 1192
QY 823 AACAGGAAGAGAGCTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGGTTT 882
DB 1193 AACAGGAAGAGAGCTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGGTTT 1252
QY 883 TGAATCTTCTCAATCTCATCTTCTTCTTCCGCTCTTTCTTTTCCAAAGGTAATAGGAAC 942
DB 1253 TGAATCTTCTCAATCTCA--TCTTCTTCCGCTCTTTCTTTTCCAAAGGTAATAGGAAC 1309
QY 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTGGAGATCTGG 1002
DB 1310 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTGGAGATCTGG 1369
QY 1003 AATTGCTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGATCGAT 1062
DB 1370 AATTGCTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGATCGAT 1429
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATTATAGCTACCCAGAAATTTGGCTTGACCTTGATGG 1122
DB 1430 CTAAGTTGACCGATCAGTTAGCTCGATTATAGCTACCCAGAAATTTGGCTTGACCTTGATGG 1489
QY 1123 AGAGATCCATGTTTCATGTTACTGGGAAATGATTTGTATGTGATGTAATGAAATCTGAAC 1182
DB 1490 AGAGATCCATGTTTCATGTTACTGGGAAATGATTTGTATGTGATGTAATGAAATCTGAAC 1549
QY 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1242
DB 1550 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1609
QY 1243 GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGCTTG 1302
DB 1610 GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTAGCTTG 1669
QY 1303 AACAGAAAGCTATTCTGATTTCAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTG 1362
DB 1670 AACAGAAAGCTATTCTGATTTCAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTGTG 1729
QY 1363 TGTGTCAGC 1372
DB 1730 TGTGTCAGC 1739

RESULT 4

US-09-737-626A-30
; Sequence 30, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flaisinski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; PRIOR FILING DATE: 2002-02-25
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 polynucleotides + Act2 in
; OTHER INFORMATION: tro
US-09-737-626A-30

Query Match 81.3%; Score 1174; DB 4; Length 1742;
Best Local Similarity 99.3%; Pred. No. 1.3e-308;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGATATAATTCAGTAATCGTTTGAC 223
DB 533 CAACATATTTTATGATGCAAGAGTCAGCATATGATATAATTCAGTAATCGTTTGAC 592
QY 224 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
DB 593 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 652
QY 284 ATGAAACATTTGATCTTATTTGATATAAATATCCATAAACAATCATGAAAGACATTTCTT 343
DB 653 ATGAAACATTTGATCTTATTTGATATAAATATCCATAAACAATCATGAAAGACATTTCTT 712
QY 344 TCAGGTCGTAATTAATTAATGATACAAATTTCTAATAGAAAACGAATTAATTAAGTTGAT 403
DB 713 TCACGGTCGTAATTAATTAATGATACAAATTTCTAATAGAAAACGAATTAATTAAGTTGAT 772
QY 404 TGTATGAAATCTAATTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 463
DB 773 TGTATGAAATCTAATTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 832
QY 464 CGGTTTAACTTAACCACTAAAAAAGCGAGCTGTCTATGTAACACCGCGATCGAGCAGGTC 523
DB 833 CGGTTTAACTTAACCACTAAAAAAGCGAGCTGTCTATGTAACACCGCGATCGAGCAGGTC 892
QY 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGAATTAAGTT 583
DB 893 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGAATTAAGTT 952
QY 584 TAAATAATTAAGTTAACAACGAGGGAAGA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642
DB 953 TAAATAATTAAGTTAACAACGAGGGAAGAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 1012
QY 643 TGTGTCGAAATGATTCGTGTCTGTCGATTTTAAATTAATTTTAAAGGCGGAAATAA 702
DB 1013 TGTGTCGAAATGATTCGTGTCTGTCGATTTTAAATTAATTTTAAAGGCGGAAATAA 1072
QY 703 AGTTGTAAAGAGATAAACCCTCATATAAATTCATATAATTTTCCCTCCCGCTTTGAATG 762
DB 1073 AGTTGTAAAGAGATAAACCCTCATATAAATTCATATAATTTTCCCTCCCGCTTTGAATG 1132

QY 763 TCTCGTTGTCCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 822
DB 1133 TCTCGTTGTCCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1192
QY 823 AACAAAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 882
DB 1193 AACAAAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1252
QY 883 TGAATCTTCTCCCTCAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
DB 1253 TGAATCTTCTCCCTCAATCTCA--TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1309
QY 943 TCTGGATCTACTCTTATTTTCTGGATCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1002
DB 1310 TCTGGATCTACTCTTATTTTCTGGATCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCT 1369
QY 1003 AATTCGTTTAATTTGGATCTCTGTAACCTCCACTAAATCTTTTGGTTTCTAGTAAGATCGAT 1062
DB 1370 AATTCGTTTAATTTGGATCTCTGTAACCTCCACTAAATCTTTTGGTTTCTAGTAAGATCGAT 1429
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTATAGCTACCAAGAAATTTGGCTTGACCTGATCG 1122
DB 1430 CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTATAGCTACCAAGAAATTTGGCTTGACCTGATCG 1489
QY 1123 AGAGATCCATGTTTCATGTTTACCTCGGAATGATTTGTATATGTGTAATGAAATCTGAACT 1182
DB 1490 AGAGATCCATGTTTCATGTTTACCTCGGAATGATTTGTATATGTGTAATGAAATCTGAACT 1549
QY 1183 GTTGAAGTTAGATTTGAATCTGAACTCTGTAACCTGTAATGTAATGTAATGTAATGTAAT 1242
DB 1550 GTTGAAGTTAGATTTGAATCTGAACTCTGTAACCTGTAATGTAATGTAATGTAATGTAAT 1609
QY 1243 GTTGAATGAGTTTCTGTATAGATTTCTCGAAACCTTAGGATTTCTAGTGTCTAGCTG 1302
DB 1610 GTTGAATGAGTTTCTGTATAGATTTCTCGAAACCTTAGGATTTCTAGTGTCTAGCTG 1669
QY 1303 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGATTTGAACTCTTTTGTG 1362
DB 1670 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGATTTGAACTCTTTTGTG 1729
QY 1363 TGTGTCGAGC 1372
DB 1730 TGTGTCGAGC 1739
RESULT 5
US-09-737-698B-10
; Sequence 10, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737,698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1271)
; OTHER INFORMATION: Y = t/u or c
; OTHER INFORMATION: Act8 promoter polynucleotide sequence and intron
US-09-737-698B-10
Query Match 5.2%; Score 75.4; DB 3; Length 1271;
Best Local Similarity 51.6%; Pred. No. 2.4e-10;

Matches 214; Conservative 3; Mismatches 194; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCTGACAGCCAGGTACAGTTA---TCTTTACCTGTGTGCGAAATGATTCGTG 662
DB 471 AGAGGCTGTCTGACAGCCATGTACGTTATCTTTTTCGGTATGATCGAAATGATTCGTG 530
QY 663 TCTGTGATTTTAATTAATTTTTCGAAAGCCGAAATAAAGTTGTAAGAGATAAACCCG 722
DB 531 TTTGYGAATTTAAATTAATTTTCCAAATTTGAYGACTCTAAAGAAAAAATAAGTTTTC 590
QY 723 CCTATATAAATCATATAATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCACTT 782
DB 591 GATAAACCCGCTATATAAATAGTTCACACATCGGTTTATTTCTCTCCCTCAAAGAA 650
QY 783 TCATCAGCGTTTGAATCTCCGGCAGTTGACAGAGAAGAAACAAGGAAGAGACTAAGA 842
DB 651 TGCCTCGTCTTCAGCTTCATCGCGCTTCAITTCGCCGCGATAAGAGAGAGAAGA 710
QY 843 -GAGAAAGTAAGAGATAATCCAGGAGATCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 711 GGAGAAAGAGTAGCCAGTCTTTCATCGTCGTGGTTCTTGTCTCTCCGATCTCTCGA 770
QY 902 TCTTCTCTTCGCTCTTCTTCTTCCAGGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 591 GATAAACCCGCTATATAAATAGTTCACACATCGGTTTATTTCTCTCCCTCAAAGAA 650
QY 783 TCATCAGCGTTTGAATCTCCGGCAGTTGACAGAGAAGAAACAAGGAAGAGACTAAGA 842
DB 651 TGCCTCGTCTTCAGCTTCATCGCGCTTCAITTCGCCGCGATAAGAGAGAGAAGA 710
QY 843 -GAGAAAGTAAGAGATAATCCAGGAGATCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 711 GGAGAAAGAGTAGCCAGTCTTTCATCGTCGTGGTTCTTGTCTCTCCGATCTCTCGA 770
QY 902 TCTTCTCTTCGCTCTTCTTCTTCCAGGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 591 GATAAACCCGCTATATAAATAGTTCACACATCGGTTTATTTCTCTCCCTCAAAGAA 650
QY 783 TCATCAGCGTTTGAATCTCCGGCAGTTGACAGAGAAGAAACAAGGAAGAGACTAAGA 842
DB 651 TGCCTCGTCTTCAGCTTCATCGCGCTTCAITTCGCCGCGATAAGAGAGAGAAGA 710
QY 843 -GAGAAAGTAAGAGATAATCCAGGAGATCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 711 GGAGAAAGAGTAGCCAGTCTTTCATCGTCGTGGTTCTTGTCTCTCCGATCTCTCGA 770
QY 902 TCTTCTCTTCGCTCTTCTTCTTCCAGGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 591 GATAAACCCGCTATATAAATAGTTCACACATCGGTTTATTTCTCTCCCTCAAAGAA 650

RESULT 6
US-09-737-626A-10
; Sequence 10, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Wilkinski, Stanislaw
; APPLICANT: Wilkinski, Jack
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; CURRENT APPLICATION NUMBER: US/09/737, 626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737, 626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1271)
; OTHER INFORMATION: y = t/u or c
; OTHER INFORMATION: Act8 promoter polynucleotide sequence and intron
US-09-737-626A-10

Query Match 5.2%; Score 75.4; DB 4; Length 1271;
Best Local Similarity 51.6%; Pred. No. 2.4e-10;
Matches 214; Conservative 3; Mismatches 194; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCTGACAGCCAGGTACAGTTA---TCTTTACCTGTGTGCGAAATGATTCGTG 662
DB 471 AGAGGCTGTCTGACAGCCATGTACGTTATCTTTTTCGGTATGATCGAAATGATTCGTG 530
QY 663 TCTGTGATTTTAATTAATTTTTCGAAAGCCGAAATAAAGTTGTAAGAGATAAACCCG 722
DB 531 TTTGYGAATTTAAATTAATTTTCCAAATTTGAYGACTCTAAAGAAAAAATAAGTTTTC 590
QY 723 CCTATATAAATCATATAATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCACTT 782
DB 591 GATAAACCCGCTATATAAATAGTTCACACATCGGTTTATTTCTCTCCCTCAAAGAA 650

QY 783 TCATCAGCGTTTGAATCTCCGGCAGTTGACAGAGAAGAAACAAGGAAGAGACTAAGA 842
DB 651 TGCCTCGTCTTCAGCTTCATCGCGCTTCAITTCGCCGCGATAAGAGAGAGAAGA 710
QY 843 -GAGAAAGTAAGAGATAATCCAGGAGATCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 711 GGAGAAAGAGTAGCCAGTCTTTCATCGTCGTGGTTCTTGTCTCTCCGATCTCTCGA 770
QY 902 TCTTCTCTTCGCTCTTCTTCTTCCAGGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 771 TCTTCTGTTTGTCTTTCGGATTAAGTAATTAAGAACTCCGATCTACTTGTCTTGTG 830
QY 962 CTGATCTCGATCTGTTTCTCAATTTCTTGAGATCTGGAATTCGTTTAATTT 1016
DB 831 TTGGATCTCGATTACGATTTCTAAGTTACCTTCAAAAGTTGTTTCCGATTTGATT 885

RESULT 7
US-09-737-698B-29
; Sequence 29, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Wilkinski, Jack
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737, 698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides +
; OTHER INFORMATION: tro
US-09-737-698B-29

Query Match 5.1%; Score 73.8; DB 3; Length 1800;
Best Local Similarity 51.3%; Pred. No. 7.7e-10;
Matches 213; Conservative 3; Mismatches 195; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCTGACAGCCAGGTACAGTTA---TCTTTACCTGTGTGCGAAATGATTCGTG 662
DB 1004 AGAGGCTGTCTGACAGCCATGTACGTTATCTTTTTCGGTATGATCGAAATGATTCGTG 1063
QY 663 TCTGTGATTTTAATTAATTTTTCGAAAGCCGAAATAAAGTTGTAAGAGATAAACCCG 722
DB 1064 TTTGYGAATTTAAATTAATTTTCCAAATTTGAYGACTCTAAAGAAAAAATAAGTTTTC 1123
QY 723 CCTATATAAATCATATAATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCACTT 782
DB 1124 GATAAACCCGCTATATAAATAGTTCACACATCGGTTTATTTCTCTCCCTCAAAGAA 1183
QY 783 TCATCAGCGTTTGAATCTCCGGCAGTTGACAGAGAAGAAACAAGGAAGAGACTAAGA 842
DB 1184 TGCCTCGTCTTCAGCTTCATCGCGCTTCAITTCGCCGCGATAAGAGAGAGAAGA 1243
QY 843 -GAGAAAGTAAGAGATAATCCAGGAGATCATTTCTCGGTTTGAATCTCTCAATCTCA 901
DB 1244 GGAGAAAGAGTAGCCAGATCTTTCATCGTCGTGGTTCTTGTCTCTCCGATCTCTCGA 1303
QY 902 TCTTCTCTTCGCTCTTCTTCTTCCAGGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 1304 TCTTCTGTTTGTCTTTCGATTAAGTAATTAAGAACTCCGATCTACTTGTCTTGTG 1363

QY 962 CTGGATCTCGATCTTCTGTTCTCAATTTCTCTGAGATCTGGAATTCGTTAATTT 1016
|||||
Db 1364 TTGGATCTCGATACGATTTCTTAAGATTACCTTCAAAAGTTGTTCCGATTGATT 1418
|||||

RESULT 8

US-09-737-626A-29
; Sequence 29, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flaksinski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in
; OTHER INFORMATION: tro
US-09-737-626A-29

Query Match 5.1%; Score 73.8; DB 4; Length 1800;
Best Local Similarity 51.3%; Pred. No. 7.7e-10;
Matches 213; Conservative 3; Mismatches 195; Indels 4; Gaps 2;
QY 606 AAAAGCTGTCGACAGCCAGGTCACGTTA--TCTTTACCTGGTTCGAAATGATTCGGT 662
|||||
Db 1004 AGAGGCTGTCGACAGCCATGTCACGTTATCTTTTCCGTATGATCGAAATGATTCGTC 1063
|||||
QY 663 TCTGTCGATTTAATTAATTTTTTTGAAAGCCGAAATAAAGTTGTAAGAGATAAACCG 722
|||||
Db 1064 TTGGYGAATTTAATTAATTTCCAAATTCAGACTCTAAGAGAAAAAATAGTTTTC 1123
|||||
QY 723 CCTATATAATTCATATATTTTCTCCCGCTTTGAAATGCTGTTGTCCTCCTCACT 782
|||||
Db 1124 GATAAACCCGCTATATAAATAGTTCAACACTCGGTTTATTTCTCTCCCTCAAAGAT 1183
|||||
QY 783 TCATCAGCCGTTTGTGATCTCCGGGACTTGCAGAGAGAACAGAGAGACTAAGA 842
|||||
Db 1184 TGCTCTGCTCTTCAGCTTCATCGGCCGTTGCAATTCCTCCGCGAATAGAGAGAGAA 1243
|||||
QY 843 -GAGAAAGTAAGAGATAAATCCAGGAGATTCATTCCTCGGTTTGAATCTTCTCAATCTCA 901
|||||
Db 1244 GGAGAAAGTGAAGCAGATCTTCATCGTCGGTCTTGTCTTCTCTCGATCTCTCGA 1303
|||||
QY 902 TCTTCTTCTCGCTCTTCTTCTTCCAGGTAATAGGAATCTTCTGGAATCTACTTTATTG 961
|||||
Db 1304 TCTTCTGCTTTTGTCTTTCCGATTAAGGTAATTAATAACCTCCGATCTACTTGTCTTGTG 1363
|||||
QY 962 CTGGATCTCGATCTTCTTCTCAATTTCTCTGAGATCTGGAATTCGTTAATTT 1016
|||||
Db 1364 TTGGATCTCGATACGATTTCTTAAGATTACCTTCAAAAGTTGTTCCGATTGATT 1418
|||||

RESULT 9

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants

; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match 3.7%; Score 52.8; DB 4; Length 1141;
Best Local Similarity 11.7%; Pred. No. 0.00031;
Matches 112; Conservative 359; Mismatches 471; Indels 15; Gaps 4;
QY 150 TAGCAAAATTTAAACAATTTTATGTATGCAAGAGTCACCATATGTATATTCATTC 209
|||||
Db 1002 TNNAMYYATTRWAAYAAAKARWAGNNRMVYGAAGNKGWCAAMATGMBWADTAGKVC 943
|||||
QY 210 AGAATCGTTTTGACGAGTTTCGGATCTAGTAGTAGCCATTATTATTAATGTACATACTAATCG 269
|||||
Db 942 NNNNNWTTDVRMAMKAKNNNNNAYWTACYNRAATNNKMAHKWKTHGHSKERTIH 883
|||||
QY 270 TGAATAGTGATATGATGAACAATCTGTATCTTATGTATAAATATCATATAACACATCATG 329
|||||
Db 882 HTRTCRRTKYNNNNNNARTVYVYHHAARRWMAWTRTNNNNNNNNNA-----CRNT 830
|||||
QY 330 AAAGACACTTCTTTCAGGGTCTGAATTAATTAATGATACAAATCTTAATAGAAAAAGAA 389
|||||
Db 829 RTWAABWKSHSCNNNNNNNNNNNNNTWCHYTANABBCYRANNNNNAARMARTCNNYMH 770
|||||
QY 390 AAATACGCTGTAATGTATGAAATCTAAATGAAACAAGCAACACGACGAGGACTAAACGT 449
|||||
Db 769 AAVTTHDDWCYKWNWYWDWMTTMBTTTTNNMTTSTNNNNNNNNNNNNNNNNNNNN 710
|||||
QY 450 TGCCTGGATTGACTCGGTTTAAAGTTAAACATCAACAAAAAAGGAGCTGTCATCAACACGC 509
|||||
Db 709 WKAYAHATNNWGCWNNNTDARRTNNNTVMRRRMTNTKTRWYSTTRRHHYTGATNNNN 650
|||||
QY 510 GGATCGAGGTCACAGTCATGAGCCATCAAGCAACAAAGAACTAAATCCAGGGGTGAG 569
|||||
Db 649 NNNNNNNNNNSCCTCTRMVTRVKKVWRDVTCTYVDVWADSWVWVYAN 590
|||||
QY 570 ATGATTAAATAGTTTAAATAATTTAGTTAAACAGGAGAAAGCTGTCTGACAGCAGGTCA 629
|||||
Db 589 WNRCDVYTRNTYCKSYAHSYWYNSNNMYRYRSARNSSMARWTRNNNNWWSGBVY 530
|||||
QY 630 CGTTATCTTTACCTGCTGGTCGAAATGATTCGTCGTGATTTTAAATTAATTTTGTGAA 689
|||||
Db 529 NERWAGTMMRHHNNNTDTRYYWKKWARBTFTTVYDSMCNAKSMWRGNMWRAMKMWAA 470
|||||
QY 690 AGCCGAAAAATAAGTTGTAAGAGATAAAC-----CCGCTATATAAATTCATATATTT 744
|||||
Db 469 NNDAGAMDHTVYMGNTMMERAWKNNMNAWCRRAYCCNNNNNNRACVWHKHKMRWTK 410
|||||
QY 745 CTCTCCCGCTTTGAATTTGCTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 804
|||||
Db 409 YMKACACNNNNKAMMYRVAMMYSRDNTNTDMMWTSDBWHWYTVDTYMMRANNNNN 350
|||||
QY 805 GCGGACTTGACAGAGAGAAACAAGGAAGAACTAAGAGAGA--AAGTAAGAGATAATCC 862
|||||
Db 349 NNWRBCKTTSMWMDHNMHTCTYGNNTWGSAYBMAAASMWAAAGASBNVYVWVWRMTYM 290
|||||
QY 863 AGGAGATTCATTCCTCGGTTTGAATCTCTCAATCTCATCTCTCTCTCTCTCTCTCTCT 922
|||||
Db 289 GKTMTNNNNNNKAYRYRTKTVAVACNNRYYVDVAVTMBKNNYKYCYAYBYWYBMVKGKHW 230
|||||


```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17067
; LENGTH: 95255
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(95255)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17067

Query Match          3.3%; Score 48.2; DB 4; Length 95255;
Best Local Similarity 48.6%; Pred. No. 0.042;
Matches 160; Conservative 0; Mismatches 168; Indels 1; Gaps 1;

QY 90 TGGCTACACTCGATGTAATTCGTAATCTCAACTCAACATTATATCTATATACCAACATTAGT 149
Db 42648 TGTATATATATATATATATTCGTTATATTTTGTATATATATATATATATCTGTATATATAT 42589

QY 150 TAGCAAAATTTAAACAACTATTTTATGTATGCAAGAGTCAGCATATGATATATATGATTC 209
Db 42588 TTGTATATATATATATATTCCTGTTATAT-TATATTTTGTATATATATATATATATCTGTT 42530

QY 210 AGAATCGTTTTCACGAGTTCCGATGAGTATGAGTATGATGATGATGATGATGATGATGATG 269
Db 42529 ATATTATATTTTGTATATATATCTGTTATATATATTTTGTATATATATATATATATATAT 42470

QY 270 TGAATAGTCATATGATGAACATGATGATCTTATGTTATATATATATATATATATATATATAT 329
Db 42469 TTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 42410

QY 330 AAAGACACTTCTTTCAGGGTCTGAATTAATATGATACAAATCTTAATAGAAAACGAAT 389
Db 42409 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 42350

QY 390 AAATACGTTGAATGTTATGAAATCTAAT 418
Db 42349 ATATTATATATATATATGATATATATAT 42321

RESULT 13
US-09-257-770-1/c
; Sequence 1, Application US/09257770
; Patent No. 6306596
; GENERAL INFORMATION:
; APPLICANT: Lambowitz, Allen M.
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Mohr, Georg
; APPLICANT: Beall, Clifford J.
; TITLE OF INVENTION: Methods for Cleaving Single-Stranded and
; TITLE OF INVENTION: Double-Stranded DNA Substrates with Nucleotide
; FILE REFERENCE: 24671/04007
; CURRENT APPLICATION NUMBER: US/09/257,770
; CURRENT FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5203
; TYPE: DNA
; ORGANISM: S. cerevisiae
US-09-257-770-1

Query Match          3.3%; Score 47.8; DB 3; Length 5203;
Best Local Similarity 45.8%; Pred. No. 0.014;
Matches 160; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY 230 GGATGTAGTATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 289
Db 71807 GGATGAAGTTGGACCCCTCTCTCAATCACATACAACTCAAAATAGATAAATATACCTAAA 71865

QY 290 CATTGTATCTTATGATATATATATATATATATATATATATATATATATATATATATATAT 349
Db 71867 TATAAAGGTAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 71926

QY 350 TCTGAATTAATATATGATACAAATCTTAATAGAAAACGAATTAATTAATTAATTAATTAATG 409
Db 71927 GGTAAAGCAATGATCTTAAATACAAACCCGAAGAAAATAAGATAGATGATGATGATGATG 71986

QY 410 AAATCTAA 417
Db 71987 GAAATTA 71994
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 19:55:51 ; Search time 925 Seconds
(without alignments)
9584.350 Million cell updates/sec

Title: US-09-868-744B-1

Perfect score: 1444

Sequence: 1 attatgatctcaaatacatt.....gtactggaatgtagtagtcc 1444

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183.8	82.0	1220	19	US-10-688-745-13
2	1174.8	81.4	1219	9	US-09-737-626A-9
3	1174.8	81.4	1219	16	US-10-427-169-9
4	1174.8	81.4	1219	16	US-10-427-180-9
5	1174.8	81.4	1219	18	US-10-909-860-9
6	1174.8	81.4	1219	19	US-10-920-869-9
7	1174.8	81.3	1742	9	US-09-737-626A-30
8	1174	81.3	1742	16	US-10-427-169-30
9	1174	81.3	1742	16	US-10-427-180-30
10	1174	81.3	1742	18	US-10-909-860-30
11	1174	81.3	1742	19	US-10-920-869-30

12	1114.6	77.2	12304	19	US-10-473-945-5	Sequence 5, Appli
13	1104.6	76.5	1259	9	US-09-887-384A-5	Sequence 5, Appli
14	1053	72.9	1202	9	US-09-887-384A-2	Sequence 2, Appli
15	1001.8	69.4	1342	9	US-09-887-384A-6	Sequence 6, Appli
16	950.2	65.8	1285	9	US-09-887-384A-3	Sequence 3, Appli
17	784.4	54.3	910	9	US-09-887-384A-4	Sequence 4, Appli
18	732.8	50.7	853	9	US-09-887-384A-1	Sequence 1, Appli
19	144	10.0	1468	17	US-10-425-114-14693	Sequence 14693, A
20	110	7.6	573	9	US-09-770-152-258	Sequence 258, App
21	94	6.5	94	9	US-09-887-384A-14	Sequence 14, Appl
22	84.4	5.8	97	9	US-09-887-384A-12	Sequence 12, Appl
23	79	5.5	79	9	US-09-887-384A-9	Sequence 9, Appli
24	77.4	5.4	79	9	US-09-887-384A-17	Sequence 17, Appli
25	76	5.3	2000	11	US-09-938-842A-3302	Sequence 3302, Ap
26	76	5.3	2000	11	US-09-938-842A-3302	Sequence 3302, Ap
27	75.4	5.2	1271	16	US-09-737-626A-10	Sequence 10, Appl
28	75.4	5.2	1271	16	US-10-427-169-10	Sequence 10, Appl
29	75.4	5.2	1271	16	US-10-427-180-10	Sequence 10, Appl
30	75.4	5.2	1271	18	US-10-909-860-10	Sequence 10, Appl
31	75.4	5.2	1271	19	US-10-920-869-10	Sequence 10, Appl
32	73.8	5.1	1800	9	US-09-737-626A-29	Sequence 29, Appl
33	73.8	5.1	1800	16	US-10-427-169-29	Sequence 29, Appl
34	73.8	5.1	1800	16	US-10-427-180-29	Sequence 29, Appl
35	73.8	5.1	1800	18	US-10-909-860-29	Sequence 29, Appl
36	73.8	5.1	1800	19	US-10-920-869-29	Sequence 29, Appl
37	65.4	4.5	67	9	US-09-887-384A-10	Sequence 10, Appl
38	65.4	4.5	74	9	US-09-887-384A-7	Sequence 7, Appli
39	63.4	4.4	77	9	US-09-887-384A-8	Sequence 8, Appli
40	56.2	3.9	436	11	US-09-732-627A-2501	Sequence 2501, Ap
41	52.2	3.6	633	18	US-10-021-323-7149	Sequence 7149, Ap
42	51.6	3.6	580	18	US-10-767-795-4034	Sequence 4034, Ap
43	50.8	3.5	1612	9	US-09-770-445-3	Sequence 3, Appli
44	50.6	3.5	3673778	16	US-10-312-841-1	Sequence 1, Appli
45	50.4	3.5	6070	15	US-10-311-455-213	Sequence 213, App

ALIGNMENTS

RESULT 1

US-10-688-745-13
; Sequence 13, Application US/10688745
; Publication No. US20050086712A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Dhugga, Kanwarpal
; TITLE OF INVENTION: Method to Produce para-Hydroxybenzoic Acid in the Stem Tissue of
; FILE REFERENCE: a Plant
; CURRENT APPLICATION NUMBER: US/10/688,745
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 13
; LENGTH: 1220
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-688-745-13

Query Match	82.0%	Score 1183.8	DB 19	Length 1220
Best Local Similarity	99.1%	Pred. No. 3.4e-271		
Matches 1212	Conservative 0	Mismatches 7	Indels 4	Gaps 2
Qy	164	CAACTATTTTATGTCAGAGTCAGCATATGTAATTTGATTCAAGTCGTTTCAC	223	
Db	1	CAACTATTTTATGTCAGAGTCAGCATATGTAATTTGATTCAAGTCGTTTCAC	60	
Qy	224	GAGTTCGATGTAGTAGGTCATTTATTTATGTACATACTAATCGTGAATGATGATG	283	
Db	61	GAGTTCGATGTAGTAGGTCATTTATTTATGTACATACTAATCGTGAATGATGATG	120	
Qy	284	ATGAACACATTTGTTTATTGTTATTAATAATATCCATAACACATCATGAAAGACATTTCTT	343	

Db 121 ATGAAACATTGATCTTATTTGATATAATCATAAACACATCATGAAAGACACTTTCTT 180
Qy 344 TCAGGCTCGAATTAATATGATACAAATCTAATAGAAAAAGAAATTAATACGTTGAAT 403
Db 181 TCACGGTCTGAAATTAATATGATACAAATCTAATAGAAAAAGAAATTAATACGTTGAAT 240
Qy 404 TGTATGAAATCTAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTGACT 463
Db 241 TGTATGAAATCTAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTGACT 300
Qy 464 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAACACGCGATCGAGCAGGTC 523
Db 301 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAACACGCGATCGAGCAGGTC 360
Qy 524 ACAGTCATGAAGCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAAGTT 583
Db 361 ACAGTCATGAAGCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAAGTT 420
Qy 584 TAAAAATTAGTTAAACACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAGTTAAACACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 480
Qy 643 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 702
Db 481 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 540
Qy 703 AGTTGTAAGAGATAAACCCGCCATATAAATTCATATATTTTCCCTCCCGCTTTGAATG 762
Db 541 AGTTGTAAGAGATAAACCCGCCATATAAATTCATATATTTTCCCTCCCGCTTTGAATG 600
Qy 763 TCTCGTGTCTCTCCCTCACCTTTCATCAGCGTTTGAATCTCCGGGACCTTGACAGAGAG 822
Db 601 TCTCGTGTCTCTCCCTCACCTTTCATCAGCGTTTGAATCTCCGGGACCTTGACAGAGAG 660
Qy 823 AACAAAGAAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCCGTTT 882
Db 661 AACAAAGAAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCCGTTT 720
Qy 883 TGAATCTTCTCAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 942
Db 721 TGAATCTTCTCAATCTCA- - -TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 777
Qy 943 TCTGGATCTACTTTATTTCTGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTGG 1002
Db 778 TCTGGATCTACTTTATTTCTGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTGG 837
Qy 1003 AATTCGTTTAAATTTGGATCTGGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 1062
Db 838 AATTCGTTTAAATTTGGATCTGGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 897
Qy 1063 CTAAGTTGACCGATCAGTTAGTCGATTAATAGTACAGAAATTTGGCTTGACCTTGATGG 1122
Db 898 CTAAGTTGACCGATCAGTTAGTCGATTAATAGTACAGAAATTTGGCTTGACCTTGATGG 957
Qy 1123 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTAATTCGAAATCTGAACT 1182
Db 958 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTAATTCGAAATCTGAACT 1017
Qy 1183 GTTGAAGTTAGATTGAAATCTGAACACTGTCAATGTTTAGATTGAATCTGAACACTGTTTAA 1242
Db 1018 GTTGAAGTTAGATTGAAATCTGAACACTGTCAATGTTTAGATTGAATCTGAACACTGTTTAA 1077
Qy 1243 GTTGAAGTTAGATTGTTGTATAGATTTCTCGAAACCTTAGATTGTTAGTGTGCTAGCTTG 1302
Db 1078 GTTGAAGTTAGATTGTTGTATAGATTTCTCGAAACCTTAGATTGTTAGTGTGCTAGCTTG 1137
Qy 1303 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAAGTCTTTTGTG 1362
Db 1138 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAAGTCTTTTGTG 1197
Qy 1363 TGTGTCAGCTCATAAAAAATGG 1385
|||||

Db 1198 TGTGTCAGCTCATAAACCATGG 1220
RESULT 2
US-09-737-626A-9
; Sequence 9, Application US/09737626A
; Patent No. US20020144304A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiniski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. US20020144304A1el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-09-737-626A-9
Query Match 81.4%; Score 1174.8; DB 9; Length 1219;
Best Local Similarity 98.7%; Pred. No. 4.7e-269;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
Qy 164 CAACTATTTTATGATGTCAGAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTGAC 223
Db 1 CAACTATTTTATGATGTCAGAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTGAC 60
Qy 224 GAGTTTCGGATGTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
Db 61 GAGTTTCGGATGTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Qy 284 ATGAAACATTTGATCTTATTTGATATAATATCCATAAACAATCATGATGATGATGATGAT 343
Db 121 ATGAAACATTTGATCTTATTTGATATAATATCCATAAACAATCATGATGATGATGATGAT 180
Qy 344 TCAGGCTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 403
Db 181 TCAGGCTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Qy 404 TGTATGAAATCTAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTGACT 463
Db 241 TGTATGAAATCTAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTGACT 300
Qy 464 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAACACGCGATCGAGCAGGTC 523
Db 301 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAACACGCGATCGAGCAGGTC 360
Qy 524 ACAGTCATGAAGCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAAGTT 583
Db 361 ACAGTCATGAAGCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAAGTT 420
Qy 584 TAAAAATTAGTTAAACACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAGTTAAACACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 480
Qy 643 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 702
Db 481 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 540
Qy 703 AGTTGTAAGAGATAAACCCGCCATATAAATTCATATATTTTCCCTCCCGCTTTGAATG 762
Db 541 AGTTGTAAGAGATAAACCCGCCATATAAATTCATATATTTTCCCTCCCGCTTTGAATG 600
|||||


```
Db 1018 GTTGAAGTTAGATTGAATCTGAACCTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1077
QY 1243 GTTAGATGAAGTTGTGTATAGATTCTTCCGAACCTTAGAATTTAGTGTGTTAGTTG 1302
Db 1078 GTTAGATGAAGTTGTGTATAGATTCTTCCGAACCTTAGAATTTAGTGTGTTAGTTG 1137
QY 1303 AACAGAAAGCTATTCTCGATTCAATCAGGGTTTATTTGACTGTATTGAACCTCTTTTGTG 1362
Db 1138 AACAGAAAGCTATTCTCGATTCAATCAGGGTTTATTTGACTGTATTGAACCTCTTTTGTG 1197
QY 1363 TGTTCGAGCTCATAAAAAATG 1384
Db 1198 TGTTCGAGCAGACTCACCATG 1219

RESULT 4
US-10-427-180-9
; Sequence 9, Application US/10427180
; Publication No. US20030199682A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiniski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. US20030199682A1e1 Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/427,180
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: US/09/737,626A
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-10-427-180-9
Query Match 81.4%; Score 1174.8; DB 16; Length 1219;
Best Local Similarity 98.7%; Pred. No. 4.7e-269;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTTCAC 223
Db 1 CAACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTTCAC 60
QY 224 GAGTTCGGATGTAGTAGTACCCATTAATTAATGTATACATACTAATCGTGAATAGTATATG 283
Db 61 GAGTTCGGATGTAGTAGTACCCATTAATTAATGTATACATACTAATCGTGAATAGTATATG 120
QY 284 ATGAACATCTCTATTGTATTAATATCCATAACACATCATCAAGACATCTTCTT 343
Db 121 ATGAACATCTCTATTGTATTAATATCCATAACACATCATCAAGACATCTTCTT 180
QY 344 TCAGGCTCTGAATTAATATGATACAAATCTTAATAGAAAACGAATTAATTAATGTTGAAT 403
Db 181 TCACGCTCTGAATTAATATGATACAAATCTTAATAGAAAACGAATTAATTAATGTTGAAT 240
QY 404 TGTATGAATCTTAATTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 463
Db 241 TGTATGAATCTTAATTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
QY 464 CGGTTTAACTTAACCACTAAAAAAGCGAGCTCTCATGTAAACACCGGATCCGAGCAGGTC 523
Db 301 CGGTTTAACTTAACCACTAAAAAAGCGAGCTCTCATGTAAACACCGGATCCGAGCAGGTC 360
QY 524 ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATTAGTT 583
```

```
Db 361 ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGCTGAGATGATTAATTAGTT 420
QY 584 TAAAAATTAGTTTAAACACGAGGGAAAA--GCTGCTCTGACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAGTTTAAACACGAGGGAAAAAGGCTGCTGACAGCCAGGTCACGTTATCTTTACC 480
QY 643 TGTGTCGGAATGATTTCTGTCGATTTTAAATTTATTTTGAAGGCCGAAAAATAA 702
Db 481 TGTGTCGGAATGATTTCTGTCGATTTTAAATTTATTTTGAAGGCCGAAAAATAA 540
QY 703 AGTTGTAAGAGATAAAACCGCCTATATAAATTCATATATTTTCTCCCGCTTTGAATG 762
Db 541 AGTTGTAAGAGATAAAACCGCCTATATAAATTCATATATTTTCTCCCGCTTTGAATG 600
QY 763 TCTCGTTGCTCTCTCATTTCATCAGCGGTTTTGAATCTCCGGCAGCTTGACAGAGAAG 822
Db 601 TCTCGTTGCTCTCTCATTTCATCAGCGGTTTTGAATCTCCGGCAGCTTGACAGAGAAG 660
QY 823 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCTCGTTT 882
Db 661 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCTCGTTT 720
QY 883 TGAATCTTCTCAATCTCATCTTCTTCTCGCTCTTCTTCTTCCAAAGGTAATAGAACTT 942
Db 721 TGAATCTTCTCAATCTCA---TCTTCTTCCGCTCTTCTTCTTCCAAAGGTAATAGAACTT 777
QY 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTCTGAGATCTGG 1002
Db 778 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTCTGAGATCTGG 837
QY 1003 AATTCTGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAATCGAT 1062
Db 838 AATTCTGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAATCGAT 897
QY 1063 CTAAGTTGACCGATCAGTTAGTCGATTATAGCTACCAAGAAATTTGGCTTGAACCTTGATGG 1122
Db 898 CTAAGTTGACCGATCAGTTAGTCGATTATAGCTACCAAGAAATTTGGCTTGAACCTTGATGG 957
QY 1123 AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTATGTAATGAAATCTGAAC 1182
Db 958 AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTATGTAATGAAATCTGAAC 1017
QY 1183 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1242
Db 1018 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1077
QY 1243 GTTAGATGAAGTTGTGTATAGATTCTTCGAAACCTTAGGATTTGATGTCGTACGTTG 1302
Db 1078 GTTAGATGAAGTTGTGTATAGATTCTTCGAAACCTTAGGATTTGATGTCGTACGTTG 1137
QY 1303 AACAGAAAGCTATTCTGATTCATCAATCAGGGTTTTTGTACTGTATTGAATCTTTTGTG 1362
Db 1138 AACAGAAAGCTATTCTGATTCATCAATCAGGGTTTTTGTACTGTATTGAATCTTTTGTG 1197
QY 1363 TGTTCGAGCTCATAAAAAATG 1384
Db 1198 TGTTCGAGCAGACTCACCATG 1219

RESULT 5
US-10-909-860-9
; Sequence 9, Application US/10909860
; Publication No. US20050005332A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiniski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: Novel Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/909,860
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/737,626
```


QY 763 TCTCGTGTCTCTCTCACTTTGATCAGCGGTTTGAATCTCCGGCGACTTGACAGAGAAG 822
DB 541 TCTCGTGTCTCTCTCACTTTGATCAGCGGTTTGAATCTCCGGCGACTTGACAGAGAAG 600
QY 823 AACAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCAATCTCCGTTT 882
DB 601 AACAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCAATCTCCGTTT 660
QY 883 TGAATCTTCTCAATCTCACTTTCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTT 942
DB 661 TGAATCTTCTCAATCTCA ---TCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTT 717
QY 943 TCTGGATCTACTTTATTTCTGGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG 1002
DB 718 TCTGGATCTACTTTATTTCTGGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG 777
QY 1003 AATTCGTTTAAATTTGGATCTGTAACCTCCACTTAAATCTTTTGGTTTACTAGAACTCGAT 1062
DB 778 AATTCGTTTAAATTTGGATCTGTAACCTCCACTTAAATCTTTTGGTTTACTAGAACTCGAT 837
QY 1063 CTAAGTTGACCGATCAGTTAGTCTGATTAAGCTACAGAAATTTGGCTTTGACCTTGATGG 1122
DB 838 CTAAGTTGACCGATCAGTTAGTCTGATTAAGCTACAGAAATTTGGCTTTGACCTTGATGG 897
QY 1123 AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATCGAAATCTGAAT 1182
DB 898 AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATCGAAATCTGAAT 957
QY 1183 GTTGAAGTTAGATTGAATCTGAACCTGCAACTGCAATGTTAGATTGAATCTGAACCTGTTAA 1242
DB 958 GTTGAAGTTAGATTGAATCTGAACCTGCAACTGCAATGTTAGATTGAATCTGAACCTGTTAA 1017
QY 1243 GTTGAAGTTAGATTGAATCTTCCGAACTTTAGGATTTGTAGTGTGCTGAGCTTG 1302
DB 1018 GTTGAAGTTAGATTGAATCTTCCGAACTTTAGGATTTGTAGTGTGCTGAGCTTG 1077
QY 1303 AACAGAAAGCTATTTCTGATTCATCAGGGTTATTTGATGATTTGAATCTTTTGTG 1362
DB 1078 AACAGAAAGCTATTTCTGATTTCAATCAGGGTTATTTGATGATTTGAATCTTTTGTG 1137
QY 1363 TGTTCAGCTCATAAAAATGCTCAGGCTGA 1395
DB 1138 TGTTCAGCTCAGGATCCATCGATAAGCTTTA 1170

RESULT 13

US-09-887-384A-5
; Sequence 5, Application US/09887384A
; Patent No. US20020049992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887, 384A
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1259
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Actin based
; OTHER INFORMATION: promoter
US-09-887-384A-5

Query Match 76.5%; Score 1104.6; DB 9; Length 1259;
Best Local Similarity 94.0%; Pred. No. 2.4e-252;
Matches 1198; Conservative 0; Mismatches 59; Indels 18; Gaps 4;
QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGATTAATGATTCAGAAATCGTTTGTAC 223
DB 1 CAACATATTTTATGATGCAAGAGTTTCATCTTGTATATAATTGATTCAGAAATCGTTTGTAC 60

QY 224 GAGTTCGGATGTAGTAGTAGCCATTTATTTAATGTATCATTAATCTGTGAATAGTAGATG 283
DB 61 GAGTTCGGATGTAGTAGTAGCCATTTATTTAATGTATCATTAATCTGTGAATAGTAGATG 120
QY 284 ATGAACAATTTGATCTTATTTGTATTAATATCCATTAACAACATCATCAAGAGACATTTCTT 343
DB 121 ATGAACAATTTGATCTTATTTGTATTAATATCATTAACAACATCATCAAGAGACATTTCTT 180
QY 344 TCAGGGTCTGAAATTAATATGATACAATTTCTAATAGAAAAAGAAATTAATTAACGTTGAAT 403
DB 181 TCAGGGTCTGAAATTAATATGATACAATTTCTAATAGAAAAAGAAATTAATTAACGTTGAAT 240
QY 404 TGTATGAATCTAATTTGAACAAGCCAAACCAACGAGGAGCTAAGCTGCTCGATGATGACT 463
DB 241 TGTATGAATCTAATTTGAACAAGCCAAACCAACGAGGAGCTAAGCTGCTCGATGATGAA 300
QY 464 CGGTTTAAAGTTAAACCACTTAAAAAAGCGAGCTGCTAATGTAACACCGGATCCAGGAGCTC 523
DB 301 AAGTTTAAAGTTAAACCACTTAAAAAAGCGAGCTGCTAATGTAACACCGGATCCAGGAGCTC 360
QY 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTT 583
DB 361 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTT 420
QY 584 TAAAAATTTAGTTAAACAAGGAAAAA - GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642
DB 421 TAAAAATTTAGTTAAACAAGGAAAAAAGGCTGCTGACAGCCAGGTCACGTTATCTTTACC 480
QY 643 TGTGTGCAAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
DB 481 TGTGTGCAAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 703 AGTGTGAAGAGATAAACCCGCTATATAAATCATATAATTTCTCCCGCTTTTGAATG 762
DB 541 AGTGTGAAGAGATAAACCCGCTATATAAATCATATAATTTCTCCCGCTTTTGAATG 600
QY 763 TCTCGTGTCTCTCACTTTTCAATCAGCGGTTTGAATCTCCGGGAGCTTGACAGAGAAG 822
DB 601 TCTCGTGTCTCTCTCACTTTTCAATCAGCGGTTTGAATCTCCGGGAGCTTGACAGAGAAG 660
QY 823 AACAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTT 882
DB 661 AACAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTT 706
QY 883 TGAATCTTCTCAATCTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 942
DB 707 TGAATCTTCTCAATCTCA ---TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 763
QY 943 TCTGGATCTACTTTATTTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTCTGATCTGG 1002
DB 764 TCTGGATCTACTTTATTTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTCTGATCTGG 823
QY 1003 AATTCGTTTAAATTTGGATCTGGAACCTCCACTAATCTTTTGGTTTACTAGAACTCGAT 1062
DB 824 TTTTCTGTTTAAATTTGGATGGAATTTAGATCACTAAATCTTTTGGTTTACTAGAACTCGAT 883
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACAGAAATTTGCTTGAATCTGATGG 1122
DB 884 CTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACAGAAATTTGCTTGAATCTGATGG 943
QY 1123 AGAGATCCATGTTTCATGTTTACTGGGAAATGATTTGTATATGTGAATGAAATCTGAACT 1182
DB 944 AGAGATCCATGTTTCATGTTTACTGGGAAATGATTTGTATATGTGAATGAAATCTGAACT 1003
QY 1183 GTTGAAGTTAGATTGAATCTGAACTGCAATGTTAGATTGAATGAACTGTTTAA 1242
DB 1004 GTTGAAGTTAGATTGAATGAACTGCAATGTTAGATTGAATGAACTGTTTAA 1063
QY 1243 GTTGAATGAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGATG 1302
DB 1064 GTTGAATGAGTTTGTGTATAGATTTCTCGAAACCTTTAGGATTTGTAGTGTGCTGATG 1123

Matches 1215; Conservative 0; Mismatches 62; Indels 56; Gaps 12;			
QY	158	TTTAAACAATCTTTTATGTATGCAAGAGTCAGCATATGTATAATTTGATTGAGTCAATCGT	217
Db	12	TTGCATCAACATATTTTATGTATGCAAGAGTTTCATCTTGTATAATTTGATTGAGTCAATCGT	71
QY	218	TTTTCAGAGCTTCGGATGTAGTAGCCCATTTTAAATGTACATATCTAATCTGTAATAGT	277
Db	72	TTTTCAGAGCTTCGGATGTAGTAGCCCATTTTAAATGTACATATCTAATCTGTAATAGT	131
QY	278	GATATGATGAACAATGTATCTTTATGTATAAATATCCATAAACAACATCATGAAGACAC	337
Db	132	GATATGATGAACAATGTATCTTTATGTATAAATATCCATAAACAACATCATGAAGACAC	191
QY	338	TTTCTTTTCAGGCTCGAATTAATTT-----ATGATACAATCTTAATAGAAAAACGAATTA	390
Db	192	TTTCTTTTCAGGCTCGAATTAATTTTGTAGAAATGATACAAATCTTAATAGAAAAACGAATTA	251
QY	391	AATTAGCTTGAATTTGATCAAAATCTAATTTGAACAAGCCCAACACGACGAGGACTTAACGTT	450
Db	252	AATTAGCTTGAATTTGATCAAAATCTAATTTGAACAAGCCCAACACGACGAGGACTTAACGTT	311
QY	451	GCCTGGATTGACTCGGTTTAAAGTTAAACCACTAAAAAAGCGAGCTGTATGTAAACACGCG	510
Db	312	GCCTGGAAATGAAGTTTAAAGTTAAACCACTAAAAAAGCGAGCTGTATGTAAACACGCG	371
QY	511	GATCGAGCAGGTCACAGT-----CATGAAGCCATCAAAGCAAAAGAACTAAATCCAAAGGGG	565
Db	372	GATCGAGCAGGTCACAGTCAATCCATGAAGCCATCAAAGCAAAAGAACTAAATCCAAAGGGG	431
QY	566	TCAGATGATTAATTTAGTTTAAAAATTTAGTTTAAACAGAGGGAAAA--GCTGTCTGACAGCCA	624
Db	432	TCAGATGATTAATTTAGTTTAAAAATTTAGTTTAAACAGAGGGAAAAAGGCTGTCTGACAGCCA	491
QY	625	GGTCACGGTTA---TCCTTACCTGTGTGCGAAATGATTCGTGTCTGCGAATTTTAAATTT	681
Db	492	GGTCACGGTTA---TCCTTACCTGTGTGCGAAATGATTCGTGTCTGCGAATTTTAAATTT	551
QY	682	T-----TTTTGAAGCGCGAAAAATAAAGTTGTAAAGATAAACCCGCGCTATATAA	731
Db	552	TCCAAAAATTTGATTTTGAAGCGCGAAAAATAAAGTTGTAAAGATAAACCCGCGCTATATAA	611
QY	732	ATTCAAT-----ATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGTCCTCTCACTT	782
Db	612	ATTCAATCAACACTCGATATTTTCTCTCCGCTTTTGAATTTGTCTCGTTGTCCTCTCACTT	671
QY	783	TCATCAGCGTTTTCGAACTCTCGCGGACTTTCAGACAGAGAACAGGAGAGACTAAGA	842
Db	672	TCATCAGCGTTTTCGAACTCTCGCGGACTTTCAGACAGAGAACAGGAGAGACTAAGA	731
QY	843	GAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCAT	902
Db	732	GAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCA--	790
QY	903	CTTCTTTCTCCGCTCTTTCTTTTCCAAAGTAAATAGGAACCTTTC--TGGAATCTACTTTATTG	961
Db	791	--TCCTTCTCCGCTCTTTCTTTTCCAAAGTAAATAGGAACCTTTCGTTGGAATCTACTTTATTG	848
QY	962	CTGGATCTCGATCTTGTGTTTCTCAATTTTCTTGGATCTGGAATTCG--TTTAAATTTGGA	1019
Db	849	CTGGATCTCGATCTTGTGTTTCTCAATTTTCTTGGATCTGGAATTTTAAATTTGGA	908
QY	1020	CTGTGGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGATCAG	1079
Db	909	TGGAATTTAGATCACTAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGATCAG	968
QY	1080	TTAGCTCGAATTATAGCTACAGAAATTTGGCTTGAACCTTGAATGAGGA---GATCCAATGTTTC	1136
Db	969	TTAGCTCGAATTATAGCTACAGAAATTTGGCTTGAACCTTGAATGAGGAATCGATCCATGTTTC	1028
QY	1137	ATGTTACCTGGGAATGATTTGTATATGTGAATTTGAATCTGA---ACTGTTGAAGTTAGA	1194
Db	1029	ATGTTACCTGGGAATGATTTGTATATGTGAATTTGAATCTGACTACTGTTGAAGTTAGA	1088

Search completed: May 28, 2005, 00:05:31
Job time : 932 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 19:32:03 ; Search time 5118 Seconds
(without alignments)
10739.509 Million cell updates/sec

Title: US-09-868-744b-1
Perfect score: 1444
Sequence: 1 attatgatctcaatacatt.....gtactggaatggtaggatccc 1444

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	272.8	18.9	552	9 BX531590	BX531590 Arabidops
2	202.4	14.0	327	1 AV831431	AV831431 AV831431
3	166	11.5	600	1 AV831862	AV831862 AV831862
4	163	11.3	546	1 AV525527	AV525527 AV525527
5	158	10.9	396	1 AV828177	AV828177 AV828177
6	158	10.9	562	1 AV830541	AV830541 AV830541
7	158	10.9	580	1 AV831261	AV831261 AV831261
8	158	10.9	632	5 BP561060	BP561060 BP561060
9	158	10.9	667	1 AV824265	AV824265 AV824265
10	158	10.9	671	5 BP560765	BP560765 BP560765
11	156	10.8	538	1 AV526249	AV526249 AV526249
12	152	10.6	399	1 AV829675	AV829675 AV829675
13	152	10.5	207	5 BP562156	BP562156 BP562156
14	151	10.5	297	1 AV525898	AV525898 AV525898
15	150	10.4	162	8 BH901724	BH901724 SALK 0858
16	147	10.2	774	5 BU635286	BU635286 018F03 In
17	146	10.1	463	7 N65062	N65062 20102 Lambd
18	144	10.0	438	7 N38049	N38049 19276 Lambd
19	144	10.0	597	7 T46735	T46735 9998 Lambda
20	143.8	10.0	382	1 AV545656	AV545656 AV545656
21	143	9.9	352	7 Z25952	Z25952 ATTS1229 Ve
22	143	9.9	393	7 T43759	T43759 7022 Lambda
23	142	9.8	411	7 R65270	R65270 13774 Lambd
24	139.4	9.7	253	7 T20889	T20889 2897 Lambda

RESULT 1
LOCUS BX531590/c 552 bp DNA linear GSS 04-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-481G03-019880,
genomic survey sequence.
ACCESSION BX531590
VERSION BX531590.1 GI:31408720
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P., and Weissshaar, B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and
Weissshaar, B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
Weissshaar, B.
TITLE High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4
AUTHORS Rosso, M.G., Li, Y., Strizhov, N. and Weissshaar, B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer
Zuchtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene At3g18780.
Details on the protocols used for generation of the sequence are
described in References 1-3. The sequences are generated at the MPI
for Plant Breeding Research in the context of the GABI-Kat project.
GABI-Kat is part of the German Plant Genomics program designated

ALIGNMENTS

'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.

```

FEATURES
    source
        1. .552
            Location/Qualifiers
                /organism="Arabidopsis thaliana"
                /mol_type="genomic DNA"
                /strain="Columbia 0"
                /db_xref="taxon:3702"
                /clone="CG-481G03-019880"
                /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                /ecotype="Col-0"
                /note="PCR was performed on DNA from Arabidopsis thaliana
                plants (T1) which were transformed with the T-DNA from
                vector pAC161 (GenBank accession number: AJ537514). The
                lines contain one or more T-DNA insertions. The DNA
                fragment(s) resulting from the PCR were directly sequenced
                to determine the genomic sequence flanking the insertion.
                T-DNA derived sequences were removed."

ORIGIN
    Query Match      18.9%; Score 272.8; DB 9; Length 552;
    Best Local Similarity 93.9%; Pred. No. 3.3e-54;
    Matches 294; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 1126 GATCCATGTTTCATGTTACCTGGGAATGATTTGTATATCTGAATCGAATCTGAACGTGT 1185
    |||||
Db 552 GATCCATGTTTCATGTTACCTGGGAATGATTTGTATATCTGAATCGAATCTGAACGTGT 493
    |||||

QY 1186 GAAGTTAGATTGAATCTGAACACTGTCGAATGTTAGATTGAATCTGAACACTGTTTAA-GT 1244
    |||||
Db 492 GAAGTTAGATTGAATCTGAACACTGTCGAATGTTAGATTGAATCTGAACACTGTTTAAAGGT 433
    |||||

QY 1245 TAGATGAAGTTGCTGATAGATTCTCGAACTTAGGATTTGTAGTGTGCTACGTTGAA 1304
    |||||
Db 432 TAGATGAAGTTGCTGATAGATTCTCGAACTTAGGATTTGTAGTGTGCTACGTTGAA 373
    |||||

QY 1305 CAGAAAGCTATTTCGATTCATCAAGGTTTATTTGACTGTATTGAATCTTTTGTGTG 1364
    |||||
Db 372 CAGAAAGCTATTTCGATTCATCAAGGTTTATTTGCTGTATTGAACTCTTTTGGGTG 313
    |||||

QY 1365 TTTCGAGCTCATAAAAAATGGCTGAGCTGACGATATTCAACCAATCGTGTGTGACAATG 1424
    |||||
Db 312 TTTCGAGCTCATAAAAAATGGCTGAGCTGATATTCAACCAATCGTGTGTGACAATG 253
    |||||

QY 1425 GTACTGGAATGGT 1437
    |||||
Db 252 GTACCGGTATGGT 240
    |||||

RESULT 2
AV831431
LOCUS
DEFINITION
    AV831431 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-89-B08 5',
    mRNA sequence.
ACCESSION
    AV831431
VERSION
    AV831431.1 GI:19873491
KEYWORDS
    EST.
SOURCE
    Arabidopsis thaliana (thale cress)
ORGANISM
    Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 327)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

REFERENCE
    AUTHORS
        TITLE
        JOURNAL
        COMMENT

```

```

FEATURES
    source
        1. .327
            Location/Qualifiers
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /clone="RAFL09-89-B08"
                /dev_stages="plants at various developmental stages from
                germination to mature seeds"
                /lab_host="DH10B"
                /clone_lib="RAFL9"
                /note="Site 1: BamHI; Site 2: SalI; subjected to
                dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                hr) treatments"

ORIGIN
    Query Match      14.0%; Score 202.4; DB 1; Length 327;
    Best Local Similarity 94.1%; Pred. No. 1.9e-37;
    Matches 222; Conservative 0; Mismatches 11; Indels 3; Gaps 1;

QY 716 AAACCCGGCTATATAAATTCATATATTTTCCTCCCGCTTTGAATGTCTCGTTGTCCTC 775
    |||||
Db 3 AAACCCGGCTATATAAATTCATATATTTTCCTCCCGCTTTGAATGTCTCGTTGTCCTC 62
    |||||

QY 776 CTCACCTTCATCAGCCGTTTGAATCTCGGCGACTTGACAGAGACAGAGAGAGAGAG 835
    |||||
Db 63 CTCACCTTCATCAGCCGTTTGAATCTCGGCGACTTGACAGAGAGAGAGAGAGAGAG 122
    |||||

QY 836 ACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
    |||||
Db 123 ACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
    |||||

QY 896 AFTCATCTCTTCTTCCTCCGCTCTTTCTTTCCAAAGTAATAGGAACCTTCTGATCT 951
    |||||
Db 183 AFTCTCA---TCTTCTCCGCTCTTTCTTTCCAAAGCTCATATAAAATGGCTGAGGCT 235
    |||||

RESULT 3
AV831862
LOCUS
DEFINITION
    AV831862 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-96-J07 5',
    mRNA sequence.
ACCESSION
    AV831862
VERSION
    AV831862.1 GI:19873922
KEYWORDS
    EST.
SOURCE
    Arabidopsis thaliana (thale cress)
ORGANISM
    Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 600)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

```


An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda PLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

Location/Qualifiers
1..600
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF09-96-J07"
/dev_stages="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAF09"
/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) treatments"

ORIGIN

Query Match 11.5%; Score 166; DB 1; Length 600;
Best Local Similarity 93.4%; Pred. No. 1e-28; Mismatches 0; Indels 3; Gaps 1;
Matches 185; Conservative 0

QY 754 TTTGAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTG 813
|||||
Db 1 TTTGAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTG 60
QY 814 ACAGAGAACAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCAT 873
|||||
Db 61 ACAGAGAACAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCAT 120
QY 874 TCTCCGTTTGAATCTTCTCAATCTCATCTTCTTCCGCTCTTCTTCCAGGTAA 933
Db 121 TCTCCGTTTGAATCTTCTCAATCTCA---TCTTCTTCCGCTCTTCTTCCAGGTCA 177
QY 934 TAGGAACCTTCTGGATCT 951
|||
Db 178 TAAAAATGGCTGAGGCT 195

RESULT 4
AV525527
LOCUS AV525527 546 bp mRNA linear EST 18-FEB-2004
DEFINITION AV525527 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone APD2508R 5', mRNA sequence.
ACCESSION AV525527
VERSION AV525527.1 GI:8685055
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 546)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
20363093
10907847
PUBMED
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

source

Location/Qualifiers
1..546
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="APD2508R"
/tissue_type="aboveground organs"
/dev_stages="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI"

ORIGIN

Query Match 11.3%; Score 163; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 5.3e-28; Mismatches 182; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
Matches 182

QY 757 GAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTGACA 816
|||||
Db 1 GAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTGACA 60
QY 817 GAGAAGAAACAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCATCT 876
|||||
Db 61 GAGAAGAAACAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCATCT 120
QY 877 CCGTTTGAATCTTCTCAATCTCATCTTCTTCCGCTCTTCTTCCAGGTAAATAG 936
|||||
Db 121 CCGTTTGAATCTTCTCAATCTCA---TCTTCTTCCGCTCTTCTTCCAGGTCAATAA 177
QY 937 GAACCTTCTGGATCT 951
|||
Db 178 AAAATGGCTGAGGCT 192

RESULT 5

AV828177

LOCUS AV828177 396 bp mRNA linear EST 01-APR-2002
DEFINITION AV828177 RAF19 Arabidopsis thaliana cDNA clone RAF09-24-N21 5', mRNA sequence.

ACCESSION

AV828177

VERSION

AV828177.1 GI:19870237

KEYWORDS

EST.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 396)

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: mseki@rtc.riken.go.jpAn Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda PLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

Location/Qualifiers
1..396
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF09-24-N21"
/dev_stages="plants at various developmental stages from germination to mature seeds"

```

germination to mature seeds"
/lab_host="DH108"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Query Match          10.9%; Score 158; DB 1; Length 396;
Best Local Similarity 93.2%; Pred. No. 7.9e-27;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTCGTTGCTCTCTCACTTTTCATCAGCGGTTTGAATCTCCGGCGACTTCACAGAGAA 821
Db 2 GTCTCGTTGCTCTCTCACTTTTCATCAGCGGTTTGAATCTCCGGCGACTTCACAGAGAA 61
QY 822 GAACAAGGAAGAAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATCTCCGTT 881
Db 62 GAACAAGGAAGAAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATCTCCGTT 121
QY 882 TTGAATCTTCTCAATCTCATCTTCTTCTCCGCTCTTCTTCCAGGTAATAGGAAT 941
Db 122 TTGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTCTTCCAGGCTCATAAAAT 178
QY 942 TTCTGGATCT 951
Db 179 GGCTGAGGCT 188

RESULT 6
AV830541
LOCUS AV830541 562 bp mRNA linear EST 01-APR-2002
DEFINITION AV830541 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-70-008 5',
mRNA sequence.
ACCESSION AV830541.1 GI:19872601
VERSION AV830541
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 562)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
REFERENCE Large scale analysis of Arabidopsis full-length cDNA (2002b)
AUTHORS Unpublished (2002)
TITLE Plant Functional Genomics Research Group
JOURNAL RIKEN Genomic Sciences Center
COMMENT 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and Sali. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.
Location/Qualifiers
1. .562
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone_lib="RAFL9-70-008"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH108"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

FEATURES
source
1. .580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone_lib="RAFL09-87-G17"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH108"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

```

QY	762	GTCTCGTTGTCCTCTCACTTTTCATCAGCGCGTTTGAATCTCCGGCGACTTGACAGAGAA	821
Db	2	GTCTCGTTGTCCTCTCACTTTTCATCAGCGCGTTTGAATCTCCGGCGACTTGACAGAGAA	61
QY	822	GAACAGAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCACAGAGATTCATTCTCCGTT	881
Db	62	GAACAGAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCACAGAGATTCATTCTCCGTT	121
QY	882	TTGAATCTTCTCCAAATCTCATCTTCTTCTTCCGCTCTTTCTTTTCCAAGGTAATAGGAACT	941
Db	122	TTGAATCTTCTCCAAATCTCA--TCTTCTCCGCTCTTCTTCTTCCAAGCTCATAAAAAT	178
QY	942	TTCTCGATCT 951	
Db	179	GGCTGAGGCT 188	
RESULT 9	AV824265		
LOCUS	AV824265	667 bp mRNA linear EST 01-APR-2002	
DEFINITION	AV824265	RAFL6 Arabidopsis thaliana cDNA clone RAFL06-16-M11 5', mRNA sequence.	
ACCESSION	AV824265	GI:19866325	
VERSION	AV824265.1		
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 667) Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Large scale analysis of Arabidopsis full-length cDNA (2002b)		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: mseki@rtc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.		
FEATURES	Location/Qualifiers		
source	1. .667 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL06-16-M11" /dev_stage="plants at various developmental stages from germination to mature seeds" /lab_host="DH108" /clone_lib="RAFL6" /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"		
ORIGIN			
Query Match	10.9%;	Score 158; DB 1; Length 667;	
Best Local Similarity	93.2%;	Pred. No. 8.5e-27;	
Matches 177; Conservative	0; Mismatches 10; Indels 3; Gaps 1;		
QY	762	GTCTCGTTGTCCTCTCACTTTTCATCAGCGCGTTTGAATCTCCGGCGACTTGACAGAGAA	821
Db	2	GTCTCGTTGTCCTCTCACTTTTCATCAGCGCGTTTGAATCTCCGGCGACTTGACAGAGAA	61
QY	822	GAACAGAGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCACAGAGATTCATTCTCCGTT	881

```
|||||
62 GAACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGAGATTTCATTTCCGGT 121
|||||
882 TTGAATCTTCTCAATCTCATCTTCTTCTCGGCTCTTCTTCCAAAGGTAATAGAACT 941
|||||
122 TTGAATCTTCTCAATCTCA---TCTTCTCGGCTCTTCTTCCAAAGCTCATAAAAAAT 178
|||||
942 TTCTGATCT 951
|||||
179 GGCTGAGGCT 188

RESULT 10
BP560765 671 bp mRNA linear EST 20-JUN-2004
LOCUS BP560765 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-03-020 5',
DEFINITION mRNA sequence.
ACCESSION BP560765
VERSION BP560765.1 GI:48976531
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
Nakashima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
TITLE Functional annotation of a full-length Arabidopsis cDNA collection
JOURNAL Science 296 (5565), 141-145 (2002)
MEDLINE 21932900
PUBMED 11910074
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp
Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for
further details.
FEATURES
source
Location/Qualifiers
1..671
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-03-020"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/note="Site 1: SstI; Site 2: XhoI; subjected to
dehydration-treated(1,2,5,10,24 hr)"

ORIGIN
Query Match 10.9%; Score 158; DB 5; Length 671;
Best Local Similarity 93.2%; Pred. No. 8.5e-27;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTGTTCTCTCTCACTTTTCATCAGCGGTTTTTGAATCTCGGCGACTTCACAGAGAA 821
|||||
DB 2 GTCTGTTCTCTCTCACTTTTCATCAGCGGTTTTTGAATCTCGGCGACTTCACAGAGAA 61
|||||
QY 822 GAACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTTCATTTCCGGT 881
|||||
DB 62 GAACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTTCATTTCCGGT 121
|||||
QY 882 TTGAATCTTCTCAATCTCATCTTCTTCTCGGCTCTTCTTCCAAAGGTAATAGAACT 941
|||||
DB 122 TTGAATCTTCTCAATCTCA---TCTTCTCGGCTCTTCTTCCAAAGCTCATAAAAAAT 178
|||||
942 TTCTGATCT 951
```

```
|||||
179 GGCTGAGGCT 188

AV526249 538 bp mRNA linear EST 18-FEB-2004
LOCUS AV526249 Arabidopsis thaliana aboveground organs two to six-week
DEFINITION old Arabidopsis thaliana cDNA clone APZ08b09R 5', mRNA sequence.
ACCESSION AV526249
VERSION AV526249.1 GI:9685777
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 538)
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
Location/Qualifiers
1..538
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="APZ08b09R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to
six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 10.8%; Score 156; DB 1; Length 538;
Best Local Similarity 93.1%; Pred. No. 2.5e-26;
Matches 175; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 764 CTCGTTGTCTCTCCTCATCTTTTCATCAGCGGTTTTTGAATCTCCGGCGACTTCACAGAGAA 823
|||||
DB 1 CTCGTTGTCTCTCCTCATCTTTTCATCAGCGGTTTTTGAATCTCCGGCGACTTCACAGAGAA 60
|||||
QY 824 ACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTTCATTTCCGGTTTT 883
|||||
DB 61 ACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTTCATTTCCGGTTTT 120
|||||
QY 884 GAATCTTCTCAATCTCATCTTCTTCTCGGCTCTTCTTCCAAAGGTAATAGGAACTTT 943
|||||
DB 121 GAATCTTCTCAATCTCA---TCCTTCTTCGCTCTTCTTCCAAAGCTCATAAAAAATGG 177
|||||
944 CTGGATCT 951
|||||
178 CTGAGGCT 185

RESULT 12
AV829675 399 bp mRNA linear EST 01-APR-2002
LOCUS AV829675 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-49-p05 5',
DEFINITION mRNA sequence.
ACCESSION AV829675
VERSION AV829675.1 GI:19871735
```

EST.
Arabidopsis thaliana (thale cress)
SOURCE
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS 1 (bases 1 to 399)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda PLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

Location/Qualifiers
source
1..399
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-49-P05"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Query Match 10.6%; Score 153; DB 1; Length 399;
Best Local Similarity 93.0%; Pred. No. 1.2e-25;
Matches 172; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
QY 767 GTTGTCTCTCTCATCTTCATCAGCGGTTTGAATCTCCGGCGACTTGACAGAGAACA 826
Db 2 GTTGTCTCTCTCATCTTCATCAGCGGTTTGAATCTCCGGCGACTTGACAGAGAACA 61
QY 827 AGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGA 886
Db 62 AGGAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGA 121
QY 887 TCTTCTCTCATCTCATCTTCTTCTTCGCTCTTTCTTCCAGGTAATAGGAACTTCTG 946
Db 122 TCTTCTCTCATCTCACTCACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 178
QY 947 GATCT 951
Db 179 AGGCT 183

RESULT 13
BP562156
LOCUS BP562156
DEFINITION BP562156 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-56-H20 5',
mRNA sequence.
ACCESSION BP562156
VERSION BP562156.1 GI:48977922
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS 1 (bases 1 to 207)
Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,
Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M.,
Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.
TITLE Functional annotation of a full-length Arabidopsis cDNA collection
JOURNAL Science 296 (5565), 141-145 (2002)
MEDLINE 21932900
PUBMED 11910074
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
Please visit our web site (<http://pfgweb.gsc.riken.go.jp/>) for
further details.

Location/Qualifiers
source
1..207
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-56-H20"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Query Match 10.5%; Score 152; DB 5; Length 207;
Best Local Similarity 98.2%; Pred. No. 1.9e-25;
Matches 165; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 762 GTCTGTGTCT 821
Db 2 GTCTGTGTCT 61
QY 822 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTT 881
Db 62 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTT 121
QY 882 TTGAATCTTCTCAATCTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 929
Db 122 TTGAATCTTCTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 166

RESULT 14
AV525898
LOCUS AV525898
DEFINITION AV525898 Arabidopsis thaliana aboveground organs two to six-week
old Arabidopsis thaliana cDNA clone APD32b06R 5', mRNA sequence.
ACCESSION AV525898
VERSION AV525898.1 GI:8685426
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
AUTHORS A large scale analysis of cDNA in Arabidopsis thaliana: Generation
TITLE of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093

PUBMED 10907847
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 252-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 FEATURES Location/Qualifiers
 source 1..297
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="APD32b06R"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
 ORIGIN
 Query Match 10.5%; Score 151; DB 1; Length 297;
 Best Local Similarity 92.9%; Pred. No. 3.5e-25;
 Matches 170; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
 QY 769 TGTCTCTCCTCCTTCATCAGCGGTTTTCGAATCTCCGGCGACTTGACAGAGAAACACAG 828
 Db 1 TGTCTCTCCTCCTTCATCAGCGGTTTTCGAATCTCCGGCGACTTGACAGAGAAACACAG 60
 QY 829 GAAGAAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATCTCCGTTTGAATC 888
 Db 61 GAAGAAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATCTCCGTTTGAATC 120
 QY 889 TTCCTCAATCTATCTTCTTCTTCCGCTCTTCTTCCAGGTAATAGGAATTTCTGGA 948
 Db 121 TTCCTCAATCTCA---TCTTCTTCCGCTCTTCTTCCAGGCTCATAAAAAATGGCTGAG 177
 QY 949 TCT 951
 Db 178 GCT 180
 RESULT 15
 BH901724/c 162 bp DNA linear GSS 04-SEP-2002
 LOCUS SALK_085874.51.15.x Arabidopsis thaliana TDNA insertion lines
 DEFINITION Arabidopsis thaliana genomic clone SALK_085874.51.15.x, genomic survey sequence.
 ACCESSION BH901724
 VERSION BH901724.1 GI:22712605
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 162)
 Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J., and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
 Unpublished (2001)
 Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of TDNA.
 Class: TDNA tagged.

Location/Qualifiers
 1..162
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Col-0"
 /db_xref="taxon:3702"
 /clone="SALK_085874.51.15.x"
 /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
 ORIGIN
 Query Match 10.4%; Score 150; DB 8; Length 162;
 Best Local Similarity 99.4%; Pred. No. 5.5e-25;
 Matches 161; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 575 TAATTAGTTTAAAAATTAGTTAAACACAGAGGAAAAA-GCTGTCTGACAGCCAGGTACCGTT 633
 Db 162 TAATTAGTTTAAAAATTAGTTAAACACAGAGGAAAAAGGCTGTCTGACGCCAGGTACCGTT 103
 QY 634 ATCTTTACCTGTGTGTCGAAATGATTTCGTCTGTCTCGATTTTAAATTATTTTGTGAAAGGC 693
 Db 102 ATCTTTACCTGTGTGTCGAAATGATTTCGTCTGTCTCGATTTTAAATTATTTTGTGAAAGGC 43
 QY 694 CGAAAAATAAGTTGTAAGAGATAAAACCCGCCCTATATAATTC 735
 Db 42 CGAAAAATAAGTTGTAAGAGATAAAACCCGCCCTATATAATTC 1
 Search completed: May 27, 2005, 23:45:17
 Job time : 5123 secs